

RICE UNIVERSITY

**Tracking Kant's *Bête Noire*: The Significance of Hegel's Emptiness Critique for  
Contemporary Kantianism**

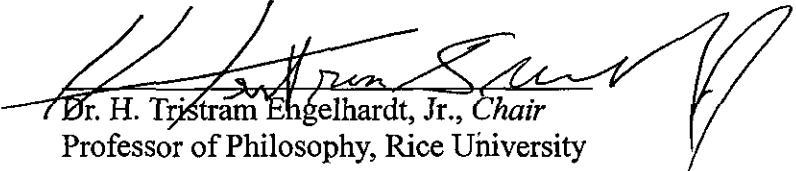
by

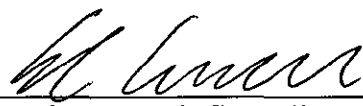
**Jesse David Perry Slavens**

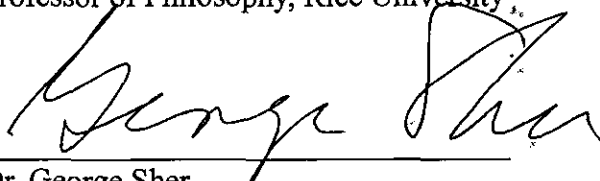
A THESIS SUBMITTED  
IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE

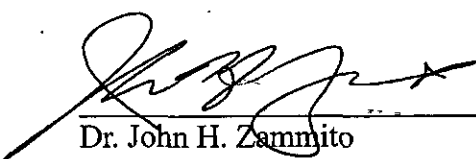
**Doctor of Philosophy**

APPROVED THESIS COMMITTEE

  
Dr. H. Tristram Engelhardt, Jr., *Chair*  
Professor of Philosophy, Rice University

  
Dr. Steven Galt Crowell  
Professor of Philosophy, Rice University

  
Dr. George Sher  
Professor of Philosophy, Rice University

  
Dr. John H. Zammito  
Professor of History, Rice University

Houston, Texas  
April, 2017

## ABSTRACT

### **Tracking Kant's *Bête Noire*: The Significance of Hegel's Emptiness Critique for Contemporary Kantianism**

by

**Jesse David Perry Slavens**

The animating contention of this dissertation is that Hegel's emptiness critique remains relevant for Kantian moral philosophy. The significance of Hegel's critique is both broad and severe. However, the relevance of Hegel's emptiness critique has been challenged on two fronts: first, that Hegel's emptiness objection misunderstands Kant's position regarding the Formula of Universal Law and the functional role of the universalization procedure; second, that Hegel's objection is limited to the Formula of Universal Law, and the content-full Formula of Humanity undermines any charges of general emptiness. Both challenges are erroneous, and the specific objective of the dissertation is to show how.

The dissertation proceeds in three parts corresponding to these rebuttals: First, in chapter 1 I provide an alternate interpretation of Hegel's emptiness objection that contextualizes the objection, thereby showing how it is an element of a larger emptiness critique. Second, in chapters 2 and 3 I consider and defend the applicability and severity of the emptiness critique for the Formula of Universal Law. Lastly, in chapters 4 and 5 I argue that Hegel's critique is not limited to the Formula of Universal Law, but can be extended to the Formula of Humanity, and is thus broadly relevant for Kantian moral philosophy. These arguments are all framed in terms of Hegel's general contention that Morality cannot furnish an immanent doctrine of duties unless empirical content is imported: when a determination or derivation of duty works, it works by importing illicit

content; when illicit content is not imported, the determination or derivation of duty fails. With the introduction of subjective, contingent, or empirical content as foundational to moral theory, however, the supremacy, objectivity, and universality of the moral theory is undone. Such content is for Kant antithetical to true morality; it is Kant's adversary and *bête noire*.

## Acknowledgments

My sincere gratitude to my adviser Dr. H. Tristram Engelhardt, Jr. whose knowledge, insight, expectations, and genuine care has made me a better scholar and person. I have been fortunate to have him as my mentor, and am honored to call him my friend.

My committee members Professors George Sher and Steven Crowell have always been generous and supportive, and their critical comments have helped me gain a better insight into my research.

I am also grateful to Dr. Clark Butler, editor of *CLIO, A Journal of Literature, History and the Philosophy of History*, and Dr. Allen Wood who both gave very helpful and extensive feedback to an earlier version of chapter one which appeared in the Fall 2014 special issue *G. W. F. Hegel*.

I would also like to sincerely thank the Philosophy Department at Rice University. Not only have faculty, students, and staff created an open, supportive, engaged, and vibrant scholarly community; the understanding and generosity of the faculty and staff has been essential to the completion of this dissertation. I am particularly and sincerely grateful for the enthusiastic and insightful conversations I have been fortunate to have with my friend Joseph Q. Adams.

Finally, I could not have done this without the enduring support of all my family, and most of all the unending, persevering support and encouragement of my wife, Suzanne Perry Slavens.

## Table of Contents

<b>Abstract.....</b>	<b>ii</b>
<b>Acknowledgments .....</b>	<b>iv</b>
<b>Introduction.....</b>	<b>vii</b>
 <b>Chapter 1: Hegel’s Emptiness Objection Reconsidered.....</b>	 <b>1</b>
I. The Standard Interpretation of Hegel’s Emptiness Objection .....	2
II. Textual Basis for an Alternate Interpretation of Hegel’s Emptiness Objection .....	7
A. Hegel’s Understanding of the Universalization Procedure.....	9
B. Heteronomy of Choice .....	13
1.Heteronomous ends and Moral subjectivity .....	15
2.Heteronomy of choice and the superfluity of the universalizability test.....	19
III. Objections .....	32
A. Triviality Objection.....	32
B. Hypothetical Universality Objection .....	36
IV. Conclusion .....	45
 <b>Chapter 2: Further Consideration of the Relevance of Hegel’s Emptiness Critique for the Formula of Universal Law .....</b>	 <b>48</b>
I. The Practical Contradiction Interpretation: Further Implications of Hegel’s Critique .....	51
A. The Problem of Natural Actions .....	52
B. Natural Actions and Moral Subjectivity .....	56
C. Conventional Actions: False-Promising .....	66
II. Maxims as Governing Principles .....	76
A. Governing Principles and Gradational Indeterminacy .....	79
B. Multiple Maxims and Moral Subjectivity .....	89
C. Ancillary Principles and the Emptiness Objection .....	94
III. Conclusion .....	98
 <b>Chapter 3: Does the Formula of Universal Law Show Too Much? .....</b>	 <b>100</b>
I. Hegel’s Excessiveness Objection .....	104
A. Superseding a Supersession: Clarifying the Structure of Hegel’s Objection .....	107
B. Objections .....	118
1.Do Hegel’s examples succeed? .....	119
2.The Logical Contradiction Interpretation and conditional maxims.....	122
3.The Practical Contradiction Interpretation .....	125
II. The ‘Murderer at the Door’ Revisited .....	130
A. Korsgaard’s Response: Antecedent Evil .....	130

B. Wood's Exegetical Rebuttal .....	135
III. Conclusion .....	150

#### **Chapter 4: Hegel's Emptiness Critique, the Formula of Humanity, and Negative Duties .....**

I. Textual Support for Hegel's Broad Critique .....	158
II. The Fullness of the Formula of Humanity .....	160
III. The Derivation of Negative Duties in <i>The Doctrine of Virtue</i> and the <i>Groundwork</i> .....	166
A. The Derivation of Negative Duties in <i>The Doctrine of Virtue</i> .....	172
1. Suicide .....	172
2. Lying (Untruthfulness) .....	180
B. The Derivation of Negative Duties in the <i>Groundwork</i> .....	189
1. Suicide .....	190
2. False-promising and deception .....	204
IV. Conclusion .....	215

#### **Chapter 5: The Formula of Humanity, Morality, and the Limits of Positive Duty .....**

I. Positive Duties in <i>The Doctrine of Virtue</i> .....	225
II. The Emptiness of the Moral Will, Moral Purity, and Subjectivism .....	233
A. Identifying the Case against Moral Purity .....	234
B. Hegel's Case against Moral Purity—the Dissemblance of Morality .....	244
1. Metaphysical hypocrisy .....	244
2. Theoretical tensions .....	254
III. The Formula of Humanity, Positive Duties, and Emptiness .....	261
A. The Expansion of Positive Duty .....	262
B. The Manifold of Duties, Conflict, and Subjectivism .....	266
1. Clarifying Hegel's objection in the context of conflicting "grounds of obligation" .....	268
2. Elaborating Hegel's objection .....	275
3. Objections .....	282
IV. Conclusion .....	295

#### **Conclusion .....**

#### **Bibliography .....**

## Introduction

Purity is essential to Kant's moral philosophy and, in many ways, remains so for Kantian moral theory. For Kant, morality is ultimately and exclusively grounded in reason, which entails the elimination of the empirical as a moral foundation. This commitment to pure or *a priori* foundations is clearly expressed in the preface to the *Groundwork of the Metaphysics of Morals*<sup>1</sup> and repeated throughout Kant's developed works in moral philosophy (*G*, 4:387–92, 411; *CPrR*, 5:70–71). This core commitment takes a variety of forms within the theory. In the opening lines of the *Groundwork* we see Kant identify the *good will* as what is good without limit. All material, external, worldly, and contingent candidates for the *good* do not have this same status. We see that even happiness is not good in itself, and that the vocation of reason becomes reason itself – reason is an end which in itself brings true satisfaction (*G*, 4:394–396). Kant proceeds from this to determine duty as that which is done without any admixture of pathological or empirical motivations (*G*, 4:400–401). Having eliminated all empirical ends as candidates for an objective end, or a ground for genuine duty, Kant identifies the form of the law itself as the “preeminent” moral good (*G*, 4:401). Finally, reason itself, being the ground of self-determination and determinate law, becomes the end of reason; reason as an end is unconditioned by any subjective end or desire, and is therefore an end *in itself*. Reason is, thus, both the foundation and final end of morality.

Kant is equally clear about the objective and necessity of a pure foundation to moral philosophy. For Kant, the purity of moral philosophy has both practical and

---

<sup>1</sup> Henceforth referred to as the *Groundwork*.

theoretical importance (G, 4:411). Practically, not only is reason itself practical (G, 4:427), a pure theory of morality provides a stable, unified, principled source for its “laws and concepts” (G, 4:411). Theoretically, a pure foundation establishes the dignity of moral theory which, in turn, establishes the supremacy of rational moral theory (G, 4:411). By grounding morality in reason alone, Kant establishes a universal moral theory that applies equally and necessarily to all rational beings (G, 4:408). Thus, the pure, *a priori* foundation of Kant’s moral theory is necessary if there is to be an objective and universal moral theory (G, 4:428).

The converse must then also hold: with the introduction of subjective, contingent, or empirical content as foundational to moral theory, the supremacy, objectivity, and universality of the moral theory is undone. Such content, which I will generally term “contingent,” is for Kant antithetical to true morality; it is Kant’s adversary and *bête noire*.

With a pure morality like Kant’s, the question naturally arises: “What can be accomplished from a foundation purged of all contingent content?” The challenge this question presents can take a variety of forms. For instance, Kant was himself in conversation with one of his earliest critics, Benjamin Constant, who questioned, by means of his famous “murderer at the door” example, the stringency of the Categorical Imperative and whether the Moral Law could distinguish what is morally required in context – within the contingent particular lives of moral agents (see chapter 3). Hegel gave a systematic and comprehensive answer to the same challenge not unlike Constant’s specific objection: pure morality – or what he terms *Moralität* – does provide a universal ground and end of moral action: the rational, free subject. However, the moral dictates of



reason are limited to the formal, abstract concepts of reason. Reason takes the universal free subject as the end of moral action and commands that, with regard to this end, the moral agent must do what is *right* and promote *the good*. In this sense, *Moralität* is empty: reason cannot guide the agent beyond its universal and objective imperatives; reason alone cannot determine an “immanent doctrine of duties.” Hegel contends, then, that *Moralität* is insufficient and internally inconsistent: it demands duty but cannot provide the resources for actualizing duty. Thus, *Moralität* must be cancelled and brought up [*aufgehoben*] to the new and complete dialectical standpoint of Ethical Life [*Sittlichkeit*].

The foregoing prelude frames the question I will consider here. I said that Kantianism maintains, in many ways, the purity of Kant's moral theory.<sup>2</sup> Almost any claim about “Kantian ethics” must be qualified; Kant's influence is extensive and the tradition in his name is diverse. Nonetheless, there is a stream of thought that continues to run through this tradition which holds that reason can be the foundation of morality, and that, specifically, the command of reason in the form of the Categorical Imperative can provide actual contextually relevant duties. A large body of Kantian literature specifically focuses on the functional usefulness of the first formulation of the Categorical Imperative, the Formula of Universal Law, and the substantive end determined by *Moralität* and represented in the second formulation, the Formula of Humanity. This common theme is and must be opposed to Hegel's contention that the duties commanded

---

<sup>2</sup>See, for example, Korsgaard (1996, xi): “Moral philosophy proceeds from what I call the standpoint of practical reason.... To respect others as ends-in-themselves is to treat them as fellow inhabitants of the standpoint of practical reason.” Herman (1993, 232; see also 238) is also representative: “Because the Moral Law applies to human beings with necessity and so independently of contingent interests, the ground of obligation must ‘be sought a priori solely in the concepts of pure reason’ (G389). Nothing in the account of moral judgment and value that I have given disregards this.”

by reason remain formal and abstract. A crucial component of this opposition is a rejection of Hegel's contention that *Moralität* is empty. One locus of this opposition has come to be known as Hegel's "emptiness objection."

It is at this critical moment within the dialectic that I pick up the story. I do so by considering interpretations and responses to Hegel's emptiness objection and rejecting two key rebuttals: first, that Hegel's emptiness objection in targeting the Formula of Universal Law misunderstands this procedural formulation of the Categorical Imperative and is, consequently, moot; second, that Hegel's emptiness objection is limited to the Formula of Universal Law, and that the Formula of Humanity, in particular, is sufficiently content-full to derive or determine the particular or contextually relevant duties of moral agents. These rebuttals are in error: Hegel understood the proper function of the Formula of Universal Law, and Hegel considered *Moralität* – represented by Kantian practical philosophy – and not merely the Formula of Universal Law to be empty. The failure to recognize and appreciate the full consequence and extent of Hegel's objection, I contend, is a product of focusing too narrowly on Hegel's contention that the Formula of Universal law is empty because it is an empty formalism, and not attending to surrounding and supporting arguments. Said differently, we must suspend a consideration of Hegel's specific "emptiness objection" until we have a more comprehensive understanding of the surrounding arguments that make up Hegel's general emptiness *critique*. It is only after Hegel's surrounding arguments are incorporated that his emptiness objection can be seen as an objection and the extent of his critique appreciated.

Explicating Hegel's emptiness critique is an extensive process. The first step is providing an alternate interpretation of Hegel's emptiness objection that incorporates

surrounding arguments. This alternate interpretation incorporates additional arguments that, one, support the emptiness contention and, two, show the broad consequences of emptiness. Furthermore, Hegel's arguments are often dense, opaque, and given without elaboration; I aim to remedy these shortcomings. Even so, Hegel's objection is, broadly speaking, straightforward. The emptiness of *Moralität* can be demonstrated in two complementary ways: when a determination or derivation of duty works, it works by importing illicit content; when illicit content is not imported, the determination or derivation of duty fails. Thus, with every attempt to determine duty by the Categorical Imperative beyond the formal determination of duty, Kantians must import illicit content in such a way that this desideratum is undermined or underdetermined. That, at least, is the contention; to what extent Hegel is right regarding the general emptiness of Kantianism will be determined by the arguments of each chapter.

The expansion of Hegel's emptiness objection to an emptiness critique takes two general forms. Regarding the incorporation of additional arguments, I argue that the emptiness objection depends not simply on the illicitness of a formal contradiction, but rather on the superfluity of the universalization procedure which produces only tautologies. I demonstrate this by arguing that the success of the universalization procedure requires the incorporation of substantive conceptual content into the will and maxim of the agent; in this way, the results of the universalization test are vacuous and the test itself is superfluous. It is only after we see the universalization procedure as superfluous that the empty formalism objection functions as an objection; this is because all that is left of the Formula of Universal law, once the illicit content is removed, is the illicitness of a contradiction.

Explicating Hegel's view regarding the incorporation of illicit content into the maxim and will of the agent requires introducing his argument for a "heteronomy of choice." On this point, Hegel contends that the agent can introduce content into his maxim in two senses: first, the agent determines the normative concepts employed in the consideration and evaluation of a maxim; second, the agent determines the intentional good, and thus essence, of his action. This heteronomous intention is integrated into his maxim. Hegel argues that this second sense of heteronomous choice entails an unacceptable subjectivism, and this subjectivism has extensive consequences for the emptiness of *Moralität*. In this way, the focus of Hegel's argument shifts: the emptiness of *Moralität* is really a byproduct of showing the inextricable role of contingent (heteronomous) content in the determination of actual duties. Hegel shows the emptiness of pure morality by revealing the substantive role of Ethical Life.

Before turning to a synopsis of the chapters, a few caveats. First, I am not concerned with defending Hegel's positive view. I make no claims and offer no defense regarding the success of a dialectical transition to Ethical Life; rather, my objective is to show how Hegel's emptiness critique reveals a deficiency to which Hegel contends Ethical Life is a solution. Second, some of the topics and arguments presented here are not new but have been considered in Anglo-American philosophy for over fifty years. Indeed, one of Hegel's central contentions, that successful universalization is contingent on the right (and objectively determined) description of actions, was perceived by Elizabeth Anscombe<sup>3</sup> and was, thereby, the catalyst for Onora O'Neill's influential treatment of the issue in both *Acting on Principle* and *Constructions of Reason*. The

---

<sup>3</sup> For a succinct example in print, see *Modern Moral Philosophy* (1958, 2).

current project is nonetheless important and unique, first, for getting Hegel's arguments right, second, for showing Hegel's systematic and extensive treatment of the issue, and third, for showing the error and limitations of Kantian rebuttals that still frame the debate – most notably, that Hegel's emptiness objection is limited to the Formula of Universal Law. Lastly, I cannot hope to cover every Kantian interpretation or argument. Kant's own arguments for the derivation or determination of duties are themselves extensive; the interpretation and variation of these arguments comprises a vast secondary literature. Accordingly, I focus on representative and influential arguments within this literature; likewise, when critiquing the derivation of duties in the *Doctrine of Virtue* I select, what I take to be, the more contentious and important duties (against suicide and lying) as heuristic examples and suggest the arguments applied to these derivations may be extended to the derivation of other duties. My hope is that this necessary circumscription nonetheless serves to demonstrate how Hegel's critique may be applied to the Formula of Universal Law as well as the Formula of Humanity.

\*\*\*

The dissertation begins with a reconsideration of Hegel's emptiness objection and what I call the "standard of interpretation" of this objection in the Kantian secondary literature. The standard interpretation holds that Hegel has grossly misunderstood Kant and the proper function of the universalization procedure. This interpretation attributes to Hegel the view that in order for the universalization procedure to work external content must first be assumed. Thus, for example, in order for the deposit maxim to generate a contradiction the validity of the institution of property must first be assumed. The

emptiness objection on this interpretation is readily rebutted by showing that the relevant contradiction generated by the universalization procedure is an internal contradiction of will. By giving a closer reading of the text, I provide an alternate interpretation of the emptiness objection where the necessary assumption for the successful function of the universalization test does not consist in assuming the external validity of any institution or convention, but rather in the internal normative conceptual content incorporated into the agent's will and maxim. The universalization procedure is employed in terms of the normative content posited by the agent in the formulation of his maxim. This content is, first, a product of a "heteronomy of choice," which in this context simply means that normative content – the thick concept of, say, "property" – used by the agent is a product of the agent's arbitrary will (*Willkühr*), and therefore contingent; this contingency means the universalization of the maxim could be in terms of a different normative specification. Both the interpretation and the consequences of this alternate interpretation are supported by Hegel's counter-factual argument: if we were to posit a different normative specification, the result of the universalization test would conform to this new normative specification. Hegel concludes from this that the universalization procedure produces only tautologies: the proscription against keeping a deposit for which there is no proof is arrived at because the normative specification in terms of which the universalization procedure was employed antecedently included this proscription. In this way, the universalization test is superfluous and the Formula of Universal Law is an empty formalism.

Chapter 2 continues with the Formula of Universal Law. This continuation is necessary and enlightening for a few reasons. First, the interpretation furnished in chapter

1 only considers the application of the emptiness critique to one example, and so the general applicability of Hegel's critique still needs to be shown. The applicability of Hegel's critique is also questionable insofar as it has only been applied under a particular construal of maxims. How maxims should be formulated is a contested issue, and so Hegel's critique must be adapted to different interpretations of maxims. Accordingly, I consider Hegel's critique in terms of Korsgaard's "thick" formulation of maxims, which includes specific intentions, and Onora O'Neill's "thin" interpretation of maxims as governing principles. Considering Hegel's critique in light of these interpretations of maxims also affords the opportunity to consider Korsgaard's influential solution to the problem of natural actions and to, in response, show the insight and effectiveness of the other side of Hegel's emptiness critique – the subjectivism objection.

Chapter 3 presents the first extension of Hegel's emptiness critique beyond the standard form, which argues that the successful function of the universalization test depends on imported illicit content, by revealing the complementary aspect of the critique: if illicit content is not imported, the universalization test does not successfully function. In his presentation of the emptiness critique in the *Natural Law* treatise, Hegel briefly argues for a corollary of the emptiness objection; namely, that the abstract formalism of the universalization test will judge permissible and even dutiful actions to be impermissible. In Korsgaard's succinct phrase, the Formula of Universal Law *shows too much*. To keep the specific arguments distinct, I refer to this argument as Hegel's "excessiveness objection." The excessiveness objection as a variant of the emptiness objection is, on the one hand, not discussed at any length in the secondary literature; on the other hand, it is a variation of the problem of false-positives (the problem of the

universalization test generating a contradiction when the maxim is putatively permissible) which is a common theme in the literature on the Formula of Universal Law. The uniqueness of the excessiveness objection is twofold: first, Hegel's presentation concerns maxims of positive duty, such as a maxim of beneficence ("Help the poor"); second, in Hegel's presentation the excessiveness of the universalization test is not a product of the increased gradational specificity of maxims – the maxims that show too much are comparable to canonical examples of properly formulated maxims. Though deviating from Hegel's own examples, the excessiveness objection affords the opportunity to reconsider Constant's 'murderer at the door example.' I delve into this infamous example and rebut two very different and formidable solutions to his objection by Christine Korsgaard and Allen Wood.

In chapters 4 and 5 I extend Hegel's emptiness critique to the Formula of Humanity. This extension has particular importance. First, some Kantians argue that the Formula of Universal Law was never intended to function independently and depends on the Formula of Humanity; a version of this argument recognizes the deficiencies identified by Hegel's critique, and contends that the Formula of Humanity provides the moral content and background knowledge which successfully augments the Formula of Universal Law; lastly, the Formula of Humanity, which commands that the humanity in one's own person or that of another always be treated as an end and never as a mere means, is evidently content-full and, therefore, ostensibly immune from the charge of emptiness.

Chapter 4 directs Hegel's emptiness critique to the derivation of negative duties in both the *Doctrine of Virtue* and the *Groundwork*, respectively. The application of Hegel's



critique to the Formula of Humanity differs significantly from that to Formula of Universal Law. The emptiness of the Formula of Universal Law was argued for by demonstrating how the universalization procedure depends on the antecedent incorporation of a specific and contingent normative specification into the will of the agent. In the derivation of negative duties from the Formula of Humanity there is no contradiction generated, and, thus, there are no terms of a contradiction that could show the derived duty to be contingent. Nonetheless, I argue that the general charge of Hegel's critique holds: the derivation of particular duties requires particular, independent moral content. I focus my arguments on the duties against suicide, untruthfulness and deceit (the latter two duties correspond to the *Doctrine of Virtue* and the *Groundwork*, respectively).

The derivation of duties in the *Doctrine of Virtue* and the *Groundwork* differ. I argue, first, that the derivation of duties in the *Doctrine of Virtue* relies on a syllogistic argument form that incorporates unique intermediate premises; consequently, and contrary to Kantian interlocutors, these intermediate premises require ancillary syllogistic arguments and moral judgment that is independent from the Formula of Humanity. Next, I consider the derivation of duties in the *Groundwork* defended by both Wood and Korsgaard. These derivations are, in contrast to the derivations in the *Doctrine of Virtue*, more direct and rely on the general intermediate premise that the maxims of the proscribed actions (suicide and deception) disregard the humanity of the relevant person (oneself or another). Moreover, this intermediate premise does not rely on an ancillary syllogism. In response, I argue that in the case of suicide the relevant sense of regard is intentional, and therefore the derivation of the duty is susceptible to Hegel's subjectivism

objection. In the case of deception, I argue that the relevant sense of disregard is one of circumventing or coopting another agent's rational will. In this way, deception is not directly impermissible, but impermissible in virtue of disregard. However, such disregard – the circumventing of coopting of another's rational will – is not a sufficient condition for the impermissibility of a maxim of action. Therefore, a sufficient mediating principle with which we could derive a duty against deceit from the Formula of Humanity has not been provided. Most importantly, such a mediating principle has been shown to be necessary, and consequently, the Formula of Humanity has been shown to be insufficient for deriving critical duties and thus an insufficient ground for an immanent doctrine of duty.

Chapter 5 extends Hegel's critique to the determination of positive duties in *The Doctrine of Virtue*. This application presents the greatest challenge for extending Hegel's emptiness critique for the simple reason that Kant does not attempt to *derive* duties at all; neither does he attempt to determine particular duties. Moreover, the general duties Kant does determine – self-perfection and practical love – follow directly from the Formula of Humanity. I adapt Hegel's critique to the general determination of imperfect positive duties with three arguments. Importantly, two of these arguments draw substantially from Hegel's earlier works, including the *Phenomenology of Spirit* and *Faith and Knowledge*. This is a disputable tactic, as Hegel is not working in or utilizing the same theoretical framework in these early works as in his more mature philosophy. Nonetheless, I think the appropriation of arguments from these works is doubly warranted: first, though there are significant differences between Hegel's early and later works, certain themes are common between them; second, the arguments from Hegel's earlier works on this topic

are influential within the Kantian secondary literature, and are used, by some, in parallel with thematically similar arguments from his later works. Even so, when appropriate I corroborate the arguments from these earlier texts with passages from the *Elements of Philosophy of Right* and *The Philosophy of Mind*.

Turning to the specifics of my adaption of Hegel's critique: First, Allen Wood explicates a version of the emptiness objection which targets the emptiness of the moral will. Taking this explication as my lead, I elaborate on Hegel's position and argue that Hegel presents the Kantian position with a dilemma: maintaining moral purity, which is both a positive duty to oneself and central to Kant's theory of moral worth, prevents *Moralität* from determining particular duties because the determination of particular duties necessarily involves empirical motivation; failing to determine particular duties, however, entails an unacceptable subjectivism regarding one's actual duty. This dilemma depends on the controversial premise that particular duties necessarily implicate empirical motivations. This is a premise the Kantian rejects. To support Hegel's case, I turn to the *Phenomenology of Spirit* and elaborate on his reasons for why, first, Kant is not in earnest about moral purity and, second, how Kant's theory of motivation is problematic on both metaphysical and theoretical grounds. These considerations contribute to undermining the cogency of moral purity as well as the normative weight moral purity has *qua* duty to oneself.

Second, I briefly argue that the conclusion of chapter 4 – namely, that an independent moral principle necessary for the derivation of negative duties has not been furnished – broadens the scope of permissible actions. A broadening of permissible actions (e.g., euthanasia or deceit) directly entails a broadening of possible positive

duties. Simply, if voluntary euthanasia, for example, is not proscribed, then assisting in another's voluntary euthanasia is not only permissible, it *may* be counted as realization of the duty of beneficence.

Lastly, I consider the underdetermination of positive imperfect duties in the *Doctrine of Duty* in light of Hegel's critique. I begin with Hegel's contention that from the standpoint of *Moralität* the moral agent is confronted by a "manifold of duties." I assess this claim in terms of Kant's position on the possibility of a conflict of duties, and conclude that he acknowledges, in the context of positive imperfect duties, the possibility of conflict between competing "grounds of obligation." Resolving such conflicts, according to Kant, is a matter of moral judgment and the relative "weight" of the competing grounds of obligation. Thus, it may appear that on this score Kant and Hegel agree – Hegel also thinks such judgment is left to the subject. However, Hegel perceives dire implications in the subjective determination of our *actual* duties, and argues that it is compatible with the elevation of the particular over the universal, which – for both Kant and Hegel – constitutes evil. I illustrate Hegel's objection by considering J. J. Thomson's famous example of Henry Fonda's cool healing hand and argue that it represents the possibility of a "threshold of opprobrium" beyond which failing to act beneficently would be condemnable; and yet the inextricable subjectivism of Kant's moral theory is compatible with not acting beneficently and, thus, formally compatible with opprobrium. Finally, I consider and reject an objection to Hegel's objection based in a rigoristic reading of Kantian positive duty that gives lexical priority to moral perfection and beneficence. Far from undermining Hegel's objection, I argue that a closer reading of Kant's arguments concerning positive duties to oneself and others, and their relative

weight, advances Hegel's objection and magnifies the formal compatibility of Kant's underdetermination of positive duties with the elevation of the particular over the universal and, thus, evil.

\*\*\*

Above I noted that I do not here defend the success of Hegel's positive project, the transition to Ethical Life. However, the development and extension of Hegel's emptiness critique emphasizes the deficiency of *Moralität* and this deficiency discloses the need for historically, culturally, and discursively contingent content for an immanent doctrine of duty and the actualization of the right and the good. In order to appreciate this deficiency to which Hegel thinks Ethical Life – or *Sittlichkeit* – is the answer, it may be helpful to clarify Hegel's concepts of *Moralität* and *Sittlichkeit*.

I will begin with the subject of Hegel's emptiness critique: *Moralität* or "Morality." According to Hegel, the foundational categorial moment of *Moralität* is the adoption of a universal yet anonymous standpoint with regard to the subject. From this standpoint, my own self *qua* subject is taken up as an end (i.e., "has being *for itself*," *PR*, §104). Within Morality the will of the subject, which exists in-itself in Abstract Right and must be respected, becomes the object and interest of the subject's will. Moreover, insofar as the will of others has the same "existence" as my will (*PR*, §112), the Moral point of view entails a universal standpoint with regard to subjectivity *per se*. This dialectical progress bears remarkable similarity to Kant's own argument for rational nature as a universal end-in-itself where the rational agent, who "necessarily represents his own existence" as an absolute and unconditioned end, recognizes that "every other

rational being also represents his existence in this way consequent on the same rational ground that also holds for” oneself (*G*, 4:429-430).

The adoption of the will as an end *in and for itself* has normative implications. Abstract Right required respect for the will of others – my relation to the will of others, in Hegel’s terminology, contained only “negative determinations” or prohibitions on action. Morality, in contrast, which involves the adoption of my will as an end, entails a positive “inner attitude” regarding my own will, my own freedom. However, the recognition of other wills with the same existence (a will that is *in and for itself*) further entails that the subject assumes positive determinations in relation to others: “two wills with a positive reference to one another [is] now posited” (*PR*, §112, *Addition*). Thus, in addition to right, in “the moral sphere...the welfare of others is also involved” (*ibid.*). In this way, Hegel argues – in a way which, again, resembles Kant – for welfare and duty (the good and the right) as formal and universal corollaries of Morality.

These parallels with Kant’s arguments for Morality and rational freedom (practical reason) as an *end-in-itself* present the opportunity to emphasize Hegel’s admiration and positive view of Morality. Hegel considered Kant’s moral philosophy to be the pinnacle and representative form of Morality. Additionally, and most importantly, Hegel recognized the profound contribution Kant made to moral thought specifically with regard to the realization of one’s genuine freedom in duty: “I should do my duty for its own sake, and it is in the true sense my own objectivity that I bring to the fulfillment in doing so. In doing my duty, I am with myself and free. The merit and exalted viewpoint of Kant’s moral philosophy are that it has emphasized this significance of duty” (*PR*, §133, *Addition*; see also §135 *Addition*). Thus, it is crucial to appreciate that Hegel does

not reject Morality in the sense of abolishing or eradicating Morality from our practical philosophy. Rather, the “cancellation” [*aufgehoben*] of Morality is based in Hegel’s contention that Morality (as well as Abstract Right) cannot exist in actuality *alone*: “The sphere of right and that of morality cannot exist independently [*für sich*]; they must have the ethical as their support and foundation” (*PR*, §141 *Addition*). Morality is an essential component of a complete practical philosophy realized in Ethical Life.

The limitation and incompleteness of Morality is a consequence of the merely formal determination of subjectivity and its necessary corollaries of welfare and duty. We can think of free subjectivity, duty, and universal welfare as the architecture of a complete Ethical Life. That is, these conceptual determinations are, in themselves, essential to Ethical Life, but also abstract and unable to provide particular ethical conclusions. To elaborate: duty as such can only command that we do *right*; but *acting* according to duty requires “particular content and a determinant end” (*PR*, §134). Likewise, promoting universal welfare can, from the purely Moral standpoint, only be understood as promoting the *good* for all. What the good is that we should promote, for whom we should promote it, who in particular should promote it, and which good among competing goods we should choose, all remain un- or underdetermined. Thus, doing right and promoting the good are, in themselves, empty. Hegel concludes from this critique that Morality, in itself, is “one-sided and incomplete.” This is Hegel’s terminology for those categories which imply a concept but cannot themselves comprehend the implied concept; that is, the concept cannot be realized through the current categorial resources (in this case, the resources of Morality). Or again, the category of Morality includes the concept of duty and welfare; the concept of duty includes the concept of *right action*;

however, from the standpoint of Morality alone we cannot comprehend which possible right action is my actual duty. Hegel concludes that we must move on to a more complete category.

This more complete category is *Sittlichkeit*, or Ethical Life, the most critical concept in Hegel's practical philosophy. "For the *good* as the substantial universal of freedom, but still as something *abstract*, determinations of some kind are therefore *required*" (PR, §141). *Sittlichkeit* provides the particular determinations in terms of which the universal concepts of right and duty are realized, and it is in this way that *Sittlichkeit* is the category which supports Morality; by unifying it with social, historical, and discursive content. Thus, Hegel contends that it is only from my particular social, historical, and discursive perspective that duty and welfare become concrete: from this content-full standpoint I am able to transition from the abstract determination *that* I have a duty, to the concrete determination of *this* is my duty.

A more concrete presentation of the concept of *Sittlichkeit* is in order. Concretely, *Sittlichkeit* has several aspects. It includes the social, civic, economic, and political institutions which structure one's life, as well as the inner incorporation of this institutional architecture into one's dispositions, relations with others, and self-understanding. Wood captures this second meaning concisely as "an attitude of harmonious identification" with the institutions of *Sittlichkeit* (Wood 1990, 196).

This idea of a social institution that also consists in an inner incorporation and "harmonious identification" with the institution is, perhaps, best illustrated by the social institution of the family. Family relations undoubtedly have a natural, biological foundation (I, at least, will assume this here). However, families are also conventional



phenomena: the structure, identity, nature of the various relations, inclusion and exclusion, closeness and fondness, authority and obedience, all have extraordinary variation throughout history and between cultures. Moreover, a family is not merely a formal institution – at least for most. We do not experience our family as we might view a bank or even the institution of banking (though these, too, are part of *Sittlichkeit*). We incorporate the relations and rich concepts of our familial convention into our sense of ourselves, into our identity. Family is the kind of convention that when it is lost or absent we can feel adrift or “not know who we are,” etc. This incorporation of our familial convention normalizes certain types of relations with and attitudes towards others – what is appropriate and what is obligatory is realized through this inner incorporation of the particular and contingent convention.

This sense of inner incorporation and harmonious identification must also be kept in mind when Hegel speaks of the distinctions of Ethical Life being elevated above the merely subjective and being *in and for themselves*:

[t]he objective sphere of ethics, which takes the place of the abstract good.... posits *distinctions* within itself which are thus determined by the concept. These distinctions give the ethical a fixed *content* which is necessary for itself, and whose existence is exalted above subjective opinions and preferences: they are *laws and institutions which have being in and for themselves*. (PR, §144)

The distinctions of *Sittlichkeit* are not, ordinarily, alien or given to us as options to choose from and assemble in accordance with our own preferences and opinions; we abide (we are at home) in *Sittlichkeit* such that the law, institutions, and conventions of *Sittlichkeit* have an immediate, unreflective reality and authority. In the *Addition* to §144 Hegel says that “the ethical man is unconscious of himself” and that this ethical unconsciousness is how we should understand Antigone’s proclamation “that no one knows where the laws

come from: they are eternal” (*ibid.*). The ethical man is unconscious in that he takes the contingent categorial arrangement of his Ethical Life as objectively real, not as a contingent artifice: one’s *Sittlichkeit* is, simply, “how things are.”

The harmonious incorporation of and identification with Ethical Life also means that the norms of Ethical Life are binding on us, but not as an external rule or legislation: “[a]ll of these substantial determinations are *duties* which are binding on the will of the individual” (*PR*, §148). The way in which these determinations are binding within *Sittlichkeit* is significant; they are internally binding, in that the agent adopts and incorporates these distinctions and determinations as part of his identity. “On the other hand, they are not something *alien* to the subject. On the contrary, the subject bears *spiritual witness* to them as to *its own essence*, in which it has its *self-awareness* and lives as in its element which is not distinct from itself” (*PR*, §147). One’s duty is integral to the elements of Ethical Life; and these elements are in harmonious relation with the agent himself; thus, one’s duty is an obligation but also perceived and experienced, “unconsciously,” as essential to one’s self-realization. Doing my actual duty is just *who I am*.<sup>4</sup>

*Sittlichkeit*, then, provides the solution to the problem of abstract indeterminacy and emptiness Hegel identifies in Morality. The determinate end which an agent adopts is

---

<sup>4</sup>Though given as commentary to Hegel’s earlier work, *Faith and Knowledge*, and in reference to an ancestral concept of *Sittlichkeit*, a “genuine ethic,” Cerf and Harris (1977, 184 n. 67) give a very helpful example which illuminates this concept of the unconscious ethical man: “Thus for example, any native Vermonter, hearing the sound of the fire siren or seeing suspicious smoke, will at once interrupt whatever he is doing, fill his car with the proper utensils and rush to help ‘the neighbor’ (which is a good word for what the parts of the *Volk* are to each other). The ethical character of his actions does not lie in his awareness of doing his duty; his freedom does not lie in his choosing between several equally open alternatives. His ethical character and his freedom, consist in his being one with the basic persuasion of his *Volk* concerning what to do in situations of this sort.”

the determinate end which is intelligible or coherent within and according to the content framing a particular *Sittlichkeit*. The normative specification in terms of which the agent's action is made intelligible consists in the norms and moral concepts of the *Sittlichkeit* the agent has incorporated into his own self-awareness and identity. *Sittlichkeit* is thus the shared and internalized form of life which allows us to explain to ourselves and others what we ought to do. *Sittlichkeit* provides the content for the reasons we give for our actions.

If Ethical Life is the foundation that supports and completes Morality, the question naturally arises, "Which *Sittlichkeit* should we have?" A full answer to this question takes us far beyond the scope of this dissertation. However, it is worth noting that Hegel is not arguing for and is not committed to adopting just any particular *Sittlichkeit*. In other words, Hegel's transition from Morality to Ethical Life shifts the locus of ethical judgment from the individual subject to the ethical life the subject inhabits. The necessary role of *Sittlichkeit* as the ground for particular duties and our reasons for action would seem to commit Hegel to a form of relativism: Whether a subject has done his duty, has acted in accordance with the right, has promoted welfare, or is virtuous, is necessarily determined by reference to the subject's community, the subject's *Gemeinwesen* (see *PR*, §150). It may appear, then, that particular duty is relativized to a particular *Sittlichkeit* such that contrary actions may be judged equally dutiful in a different *Sittlichkeit*.

I cannot here give an answer as to how Hegel avoids such relativism; presently, I can only state that neither Hegel nor the standpoint of Ethical Life is indifferent concerning contrary or incommensurable duties. It must be remembered that Morality has

not been eradicated, but brought into [*aufgehoben*] Ethical Life, and because of this, genuine freedom of the subject remains the end of practical philosophy. The crucial difference for Hegel is that he is concerned with *actual* freedom – that is, how freedom is realized within Ethical Life. In this way, judgment concerning incommensurable duties is made in terms of the realization of freedom. “In a state which is really articulated rationally all the laws and organizations are nothing but a realization of freedom in its essential characteristics” (*Lectures on Fine Arts*; cited in Pippin 2008, 7). Furthermore, insofar as the laws and organizations of Ethical Life are internalized by the agent, and binding on him in the form of duty, the extent to which the laws and organizations of a rational state realize freedom, is also the extent to which the duty of an agent in that state is liberating (see *PR*, §149). Moreover, the converse is also true: In a state which is articulated *irrationally*, the laws and institutions which externally constitute a *Sittlichkeit*, and which are internally binding on the subject in the form of duty, can be judged to be one-sided and incomplete in that they are *not* the “realization of freedom in its essential characteristics.” In this way, Hegel avoids unqualified relativization among communities and their *Sittlichkeiten*. The realization of freedom, *actual* freedom, provides the ground for a rational adjudication between *Sittlichkeiten*, and the framework for an ethics of liberating duty.

Transition to Ethical Life, however, is only necessary if Morality is empty and cannot provide an immanent doctrine of duty. I turn now to my contribution affirming this charge.

## Chapter 1: Hegel's Emptiness Objection Reconsidered

This chapter develops an alternative interpretation of Hegel's emptiness objection to Kant's Formula of Universal Law and universalization procedure, and situates this interpretation within a broader emptiness critique Hegel levels against Kantian Morality.<sup>1</sup> In subsequent chapters I will argue for the broad relevance of the emptiness critique for the Formula of Universal Law (chapters 2 and 3) and for extending this critique to other formulations of the Categorical Imperative (*viz.*, the Formula of Humanity; chapters 4 and 5). In this chapter my thesis is more modest and focused: to clarify what I contend is a charitable and textually supported interpretation of the emptiness charge that remains relevant and critical for contemporary Kantian moral theory.

What is called Hegel's "emptiness objection" is widely discussed within contemporary Kantian literature, and is compared to what are taken to be similar objections raised by J. S. Mill and Gottlob Tittel.<sup>2</sup> The emptiness objection to the universalization procedure is often considered in isolation. I argue this is a mistake, and

---

<sup>1</sup> I will use "Morality" capitalized to refer to what Hegel's designates *Moralität*; that is morality which understands the person *qua* subject (understood as a will "*infinite not only in itself but also for itself*"), and the subject as the locus of right and duty. (*PR*, §105; see also §107).

<sup>2</sup> Hegel's emptiness charge is regularly compared to J. S. Mill, who writes:

But when he begins to deduce from this precept any of the actual duties of morality, he fails, almost grotesquely, to show that there would be any contradiction, any logical (not to say physical) impossibility, in the adoption by all rational beings of the most outrageously immoral rules of conduct. All he shows is that the *consequences* of their universal adoption would be such as no one would choose to occur.

This comparison to Mill, who clearly holds that the only contradiction which could be generated is between the consequences of universalizing impermissible maxims and the desirability for certain social institutions, further supports my contention that Hegel has been interpreted as arguing that the contradiction to be generated is external. Allen Wood also connects the emptiness charge to Gottlob Tittel's *Über Herr Kants Moralreform* (Frankfurt: Pfahler, 1786; cited in Wood 1990).

that the narrow emptiness objection is best understood as an element within a broader *emptiness critique*. Specifically, I argue that the “empty formalism” of the moral law only serves as an objection within the context of a constellation of arguments Hegel levels against the universalization procedure. It is these other arguments, I contend, which do the real work in showing that this procedure is an empty formalism; that is, its empty formalism is a problem for Kant and Kantians *only if* Hegel’s prior arguments go through. These prior arguments concern the tautological vacuity, and thus superfluity, of the universalization procedure.

I identify “heteronomy of choice” as a crucial component underlying Hegel’s critique. By identifying heteronomy of choice in this way I connect the emptiness of the universalizability principle with Hegel’s additional contention that the function of the principle relies on an unacceptable moral subjectivism. This connection shows that both the emptiness and unacceptable subjectivity of Kantian Morality are consequences of a general objection that Morality must rely on heteronomous content, a reliance which the Kantian clearly cannot accept.

## I. THE STANDARD INTERPRETATION OF HEGEL’S EMPTINESS OBJECTION

I turn first to what I call the *standard emptiness objection*. The primary source in the Kantian literature for interpreting Hegel’s objection is a passage from *On the Scientific Treatment of Natural Law*:

Kant shows by an example: I ask whether my maxim to increase my fortune by any and all safe means can hold good as a universal practical law in the case where [appropriating] a deposit entrusted to me has appeared to be such a means; the content of this law would be that “anyone may deny having received a deposit for which there is no proof.” This question is then decided by itself, “because such a principle as a law would destroy itself since the result would be that no deposits

would exist.” But where is the contradiction if there were no deposits? The non-existence of deposits would contradict other specific things, just as the possibility of deposits fits together with other necessary specific things and thereby will itself be necessary. But other ends and material grounds are not to be invoked; it is the adopting of either one specific matter or the other. For the form, however, one of the opposed specifics is just as valid as the other; each can be conceived as a quality, and this conception can be expressed as law. (*NL*, 77)<sup>3</sup>

Various Kantian commentators cite this passage as support for the standard interpretation of Hegel. H. B. Acton, in his introduction to *Natural Law*, attributes to Hegel the position that a contradiction can only be generated if the thief is “going against the institution of property” and that a contradiction must obtain between “there being deposits and there not being deposits” (Acton, 1975, 24). Allen Wood insists that Kant, contra Hegel's objection, “neither needs to show nor tries to show that the maxim would *literally* result in a self-contradiction if universally followed.” Wood goes further, arguing that Kant is not appealing to “‘other ends or material grounds’...other than those contained in my maxim itself” (Wood, 1990, 157; italics are mine).

The standard interpretation, then, attributes to Hegel the contention that, in order to generate the requisite contradiction, material grounds must be assumed which are in some sense external to the agent's will. Another way of characterizing this objection is that Hegel argues that all the universalization test has and can show is that thieving is incompatible with a convention of property and further, that the universalizability test

---

<sup>3</sup> See *CPrR* 5:27-28. This example presents an interesting case insofar as Kant identifies the maxim as “to increase my wealth by every safe means” and then applies this general maxim to a specific case and produces a *law* (“everyone may deny a deposit which no one can prove has been made”) and it is this law that is submitted for universalization. This contrasts with the example of false-promising, an example of another violation of negative duty of right, which formulates the maxim more specifically: “his maxim of action would go as follows: when I believe myself to be in need of money, I shall borrow money and promise to repay it, even though I know this will never happen” (*G*, 4:422). One might take the deposit example as supporting the formulation of maxims as governing principles, which I address in chapter 2.

cannot answer the prior question of whether or not a world without property would be a morally inferior world to a world with property.<sup>4</sup> Clearly, if we remove the institution of deposit-making from the presumed context within which the maxim of theft is adopted, then there can be no contradiction generated by annihilating deposits. In this way, Hegel is thought to understand “the example of the deposit as if it turned on the validity of the institution of private property.” (Wood, 1990, 158)

In other words, the universalizability test on this interpretation shows the impermissibility of thieving within the institution of property, but in order for this impermissibility to have any general normative force, we must first assume the moral inferiority of a world without property. But this assumption constitutes an external “material ground” to which the Kantian cannot appeal, and which the universalizability principle cannot itself establish. Wood, again, exemplifies this standard interpretation of Hegel. “He [Hegel] appears to think that the FUL [Formula of Universal Law] is applicable to cases like the deposit, only if it can be shown that the absence of this institution would result in a contradiction” (Wood 1990, 158).<sup>5</sup>

The Kantian would find this characterization of the universalization test unsatisfactory. The Kantian desideratum is not merely to show that thieving is wrong *within* a system of property, but rather, that there is something universally wrong or

---

<sup>4</sup> This characterization is adapted from W. H. Walsh’s discussion of Hegel’s objection, cited in Acton (1975,24).

<sup>5</sup> Further support for the standard interpretation and an external contradiction is found in Acton’s (1970, 24-25) *Kant’s Moral Philosophy*:

...Hegel says that all Kant’s arguments shows is that a system without deposits is contradicted by a system with deposits. Kant makes there seem to be a contradiction in a system without deposits because he assumes that everyone would want there to be deposits, and this, says Hegel, shows that Kant was assuming the system of property and was arguing that if everyone kept what belongs to others then there would be no system of property. (cited in Korsgaard, 1996, 86)



impermissible with the thieving maxim regardless of the existence or non-existence of property. However, in order to demonstrate this, the Kantian clearly cannot assume the validity of property.

The Kantian rebuttal to the standard interpretation of Hegel's emptiness charge is that Hegel has failed to understand the intended function of the universalization test. The contradiction the test is designed to generate in the case of an impermissible maxim is *not* literal. Nor does it rely on any assumption regarding the external validity of any particular institution. The intended "contradiction lies not in envisioning a society in which there are no deposits, but in envisioning a society in which the agent and others with his purpose are making use of the deposit system even though there is no such thing. The contradiction is generated when the agent tries to will his maxim and the universalization of his maxim *at the same time*" (Korsgaard, 1996, 86). The contradiction is a practical contradiction of both willing and not willing the same state of affairs. That is, the agent undermines the means to his end by *willing* that there is property (in taking a deposit) and *willing* that there is no property (in universalizing his maxim). The validity of property (or any other normative convention or practice external to the agent's will) is entirely irrelevant insofar as the desired contradiction is generated when the agent, in the universalization of his maxim, wills that property does not exist, and insofar as he wills there to be a deposit which he can appropriate, simultaneously wills that property does exist. In this way, the relevant contradiction is entirely *internal* to the agent's own will and results in a *practical* impossibility insofar as the agent cannot realize his proposed end. The person who tries to will the universalization of this maxim is not only thereby willing a situation in which practices like deposits and promises do not exist. He is also

willing that they do exist, precisely because he is willing to *use* them to achieve his ends (Korsgaard, 1996, 95).<sup>6</sup> In this way, the Kantian rebuttal concludes, Hegel has fundamentally confused the locus of the contradiction Kant is attempting to derive, and thus the function and consequent generalizability of the universalizability test. On this point, Wood concludes “[Hegel is] so far from a correct understanding of the universal law test that he cannot be credited with having identified any of the real difficulties with it” (Wood, 1996, 161).

Because the locus of the contradiction is entirely internal to the will of the agent himself, Kant need not presuppose the validity of an institution of deposits existing independently of his will. Furthermore, the derived impermissibility should not be understood in terms of thieving *per se*. Rather, the universalizability principle reveals a contradictory structure to the agent’s will which in turn identifies a generalizable violation of duty; *viz.*, the exploitation of any convention or practice in such a way that the agent makes an exception for himself. In other words, what is wrong in the example of keeping deposit is not thieving, but taking advantage of another’s trust and in doing so exempting oneself, and only oneself, from the universal normative expectations of trust (see Wood, 1996, 158).

There is much that standard interpretation of Hegel’s emptiness objection gets right. Specifically, Kant need not appeal to any content or material grounds that are not provided by the agent’s own maxim. The contradiction to be generated is rightly understood as an internal and practical contradiction of will. However, this correct

---

<sup>6</sup> Wood’s analysis of Hegel’s objection and rebuttal is virtually identical to Korsgaard’s. “The argument shows that I will an impossible or self-contradictory world in the sense that I will both (1) that I profit by appropriating the deposit and (2) that circumstances be such that I could never gain money in this way” (Wood 1990, 157).

interpretation of Kant does not exculpate the Formula of Universal Law and universalization test from Hegel's full critique. To the contrary, ample textual evidence supports a different interpretation of Hegel's emptiness objection which effectively shows that the generation of a contradiction of will using the universalization principle is a vacuous achievement insofar as it relies on contingent normative presuppositions incorporated into the agent's will itself. This interpretation shows that Hegel was concerned with prior *conceptual assumptions* incorporated into the agent's will and maxim, and the attribution of necessity to, what are in fact, particular and contingent conceptual specifications. Moreover, this vacuity is generalizable. That is, Hegel is not simply arguing that the Kantian must presuppose the validity of, say, property; he is, rather, arguing that for any maxim, embedded in any convention, the success of the universalization test depends on the particular presupposed normative specification that already includes the relevant proscription which the test is itself designed to derive. In a word, Hegel's full critical approach first shows that the universalization procedure is *superfluous*, and that the real work lies in justification of its presupposed norm. It is only after this superfluity is established that the empty formalism of the universalizability principle itself can be appreciated *qua* objection.

## II. TEXTUAL BASIS FOR AN ALTERNATIVE INTERPRETATION OF HEGEL'S EMPTINESS OBJECTION

As other commentators of Hegel have noted, there is no decisive textual basis for attributing to Hegel the standard interpretation of the emptiness objection.<sup>7</sup> Hegel's

---

<sup>7</sup> See Sedgwick 2011

emptiness objection *may* be read *either* in terms of an external contradiction between universalization of a maxim and a presupposed normative institution *or* as pertaining to an internal contradiction of the agent's will with itself. While this interpretive equivocacy may be sufficient for showing the relevance of the emptiness objection to internal contradictions of will, there is clear textual evidence for an alternative interpretation of Hegel which shows the vacuity of a contradiction internal to the will to be the target of his objection. I develop this interpretation primarily by providing a more thorough exegesis of Hegel's key texts from *Natural Law*, supplemented by his more mature work in ethics, *Elements of the Philosophy of Right*. Central to my interpretation are often neglected passages surrounding the oft cited passage from *Natural Law* quoted in Section I and used ubiquitously to support the standard interpretation. These successive passages are conspicuously absent from many Kantian commentaries – an omission that, as will become clear, precludes a more tenable reading of Hegel. Hegel's own argumentation in these passages, however, is sparse and at times unclear; at those points, elaboration is needed. Specifically, Hegel argues that the content used in the universalization test and the subsequent legislation of pure practical reason are tautological. Secondly, Hegel contends counterfactually that *if* we were to posit opposite normative specifications, then the universalizability test would fail to generate the desired contradiction. Consideration of these further passages militates against the standard interpretation of Hegel's critique, and situating the emptiness objection in the context of supplementary arguments renders the Kantian rebuttal ineffectual.

### A. Hegel's Understanding of the Universalization Procedure

Before turning to these supplementary arguments, we must rectify the charge that Hegel has misunderstood the function of the universalization procedure and the role of maxims in that procedure. In order to appreciate Hegel's view of the universalizability test, and his understanding of how pure practical reason cannot alone provide particular duties, it is helpful to consider first what he says earlier in the *Natural Law* essay about the "empty formalism" of reason generally:

... the production of tautologies is in truth what the sublime lawgiving power of pure practical reason's autonomy in legislating consists of. The pure identity of the Understanding, expressed in theoretical terms as the principle of contradiction, remains exactly the same when it is turned to the practical form. (*NL*, 76)

Of foremost importance in this passage is Hegel's view that pure practical reason produces *tautologies*; this is the first indication that internal content, and not external content, plays a crucial role in Hegel's charge that the universalizability test of any particular maxim of conduct is empty. Additionally, Hegel's appropriation of Kant's own view regarding the "empty" or pure formalism of the Kantian faculty of the Understanding is exegetically significant. Hegel draws an analogy between Kant's dual faculty theory of cognition and the universalization principle in order to show the common formal emptiness of the Understanding and pure practical reason. Drawing this analogy between Kant's theory of cognition and the universalization principle supports the interpretation I provide below, and specifically my contention – contra Kantian interpreters – that Hegel was well aware of how the universalization test is designed to function as a test of independently provided content.

Kant's theory of cognition argues that two faculties are necessary for a complete cognition: formal understanding and sensibility. The faculty of understanding provides

formal categories of relation which grant sensible content intelligibility and thereby ground the very possibility of experience. The understanding itself is nothing but these categories of formal relation, which leads to Kant's pithy dictum that "Thoughts without content are empty, intuitions without concepts are blind" (*CPR*, A51/B75). A complete cognition, then, consists in the categories of the understanding applied the content supplied by sensibility. Hegel cites Kant's model of cognition because he thinks that "In saying this [about cognition], Kant is pronouncing judgment on the principle of duty and right set up by practical reason" (*NL*, 76). In drawing an analogy between this dual faculty model of cognition and the derivation of particular duties via the universalizability test, we see first that Hegel appreciates that the universalizability principle constitutes the formal component of the universalization procedure, and that content must be adopted from an external source into the will for the procedure to function. I will show that Hegel recognizes that this content is provided by the agent's own chosen maxim.

This analogy also provides us with insight in to the structure of Hegel's critique. Kant recognized the absurdity of trying to form a criterion of truth for particular content where this criterion is itself devoid of this content. In other words, Kant held that any meaningful criterion of truth will have to be in terms of the content it judges to be true. Hegel's critique of the pure formalism of practical reason parallels this worry. Hegel argues that "duty" and "right" – understood as more than formal categories – must be understood in terms of a given set of norms, or what he refers to as a "specification." It is only after a given normative specification has been posited that the universalization test can function successfully. In this way the universalization procedure also has a dual

structure. However, the analogy between a complete cognition and determining what we ought to do ultimately breaks down within Kant's own framework, according to Hegel. The content provided for the procedure to function plays a crucial but ultimately illicit role in the determination of one's actual duty. Indeed, Hegel will argue that the posited content is of such normative specificity that the duties derived via the universalization procedure are tautologous. Thus Hegel's critique of Kant's universalizability test amounts to saying that the important ethical work consists in independently evaluating the relevant norms by which one's duty is to be determined. This is the hard question, and once this question is settled, the universalization test is no longer needed.

Before getting to these arguments, however, we must further establish that Hegel's critique applies to internal contradictions of will with which Kant is concerned. In the surrounding passages in *Natural Law* Hegel explicitly recognizes, first, that pure practical reason is purely formal, and second, that the universalization procedure functions only by "taking up" the maxim of the agent's will into the form of the concept—that is, by universalizing content that is internal to the agent's particular will.

Kant, the man who has expounded this abstraction of the concept in its absolute purity, recognizes full well that practical reason totally renounces the content of law and can do nothing beyond making the *form* or *fitness* of the will's (*Willkühr*) maxim into supreme law. The maxim of the arbitrary will (*Willkühr*) in choosing has a content and includes a specific action, but the pure will (*Wille*) is free from specification....If the specification [given in the maxim and chosen by *Willkühr*] can be taken up into the form of the pure Concept, if it is not cancelled thereby, then it is justified and has itself become absolute through negative absoluteness. (*NL*, 75)

In a subsequent passage, Hegel continues to characterize the function of the universalization procedure in terms compatible with an internal contradiction of will:

If this formalism is to be able to promulgate a law, some matter, something specific, must be posited to constitute the content of the law. And the form given

to this specific matter is unity or universality. “That a maxim of thy will shall count at the same time as a principle of universal legislation” – this basic law of pure practical reason expresses the fact that something specific, constituting the *content of the maxim of the particular will*, shall be posited as concept, as universal. (Hegel 1976, 77; my italics)

I quote these passages to emphasize two things: First, on a charitable interpretation we cannot understand Hegel’s emptiness objection as simply insisting that the empty formalism of the universalizability principle is insufficient to provide us with particular duties, and that content external to the agent’s own will must first be assumed as valid. If the formal emptiness of the universalization test does constitute an objection – and I aim to show how it does – then that principle must be understood within a broader context. Second, Hegel is fully aware that the particular content to be tested, to be “taken up” into the universal concept, is furnished by the agent’s own maxim. If this is right, if Hegel is aware of the origin of that content in the agent’s act of volition, we must make sense of Hegel’s objection in a way accordant with his recognition that the only material grounds used in the universalization procedure are those furnished by the agent’s own maxim. We must reframe Hegel’s objection in terms of an internal contradiction of will.

If Hegel recognizes that the content tested by the universalization procedure is that of the agent’s own maxim, how can we make sense of his contention that the universalization procedure must appeal to “other material grounds?” How can the material grounds used in the universalization procedure be both internal to the agent’s will and yet illicit? Hegel’s answer is that the agent’s maxim itself incorporates illicit normative content. Hegel argues that a particular specification of, say, “property”—which may or may not be the institutions of some society—is posited prior to the universalization procedure, and this assumed normative content constitutes the ground in



terms of which the agent's action and maxim, and eventually his duty, is understood.

Hegel will argue further that if this specification of property as a moral norm is not assumed, then we cannot derive the relevant and desired particular duty to respect property. He will then conclude that the universalization procedure in deriving positive duties is superfluous once the morally right normative specification has been assumed.

Hegel's initial move in this critique of the universalization principle is to identify the heteronomous origin of the normative content incorporated into the agent's will. Accordingly, I will first establish what Hegel means by "heteronomy of choice" and then connect this concept to Hegel's additional objection, developed in the *Philosophy of Right*, concerning the unacceptable moral subjectivity of Kantian duty. I then return to Hegel's emptiness charge and argue that his arguments in this regard parallel his later argument in the *Philosophy of Right* against subjective morality.

## **B. Heteronomy of Choice**

Hegel thinks that "to introduce a content is to establish a heteronomy of choice" (NL, 76). The choice being made concerns the action that the agent performs. "The maxim of the arbitrary will (*Willkühr*) in choosing has a content and includes a specific action, but the pure will (*Wille*) is free from specification" (NL, 75). When the agent wills his maxim, and that his maxim be universalized, he introduces – through a heteronomous faculty – a content which characterizes his action. The agent *wills what to do* not merely in the trivial sense of choosing one possible action instead of another, but in the more substantive sense of willing a content, that is, particular normative concepts and ends, in terms of which his action is made intelligible.

Kant's conception of practical reason includes a containment relation which holds that a rational agent, in willing an end, necessarily also wills the means to that end.<sup>8</sup> Hegel assigns heteronomy of choice to both components of Kantian action: the end which the agent wills and the way in which he makes intelligible the means to that end. Hegel discusses the former component in the *Philosophy of Right*, while the latter is the focus of *Natural Law*. In the *Philosophy of Right* Hegel argues that the end which the agent adopts is contingent, and that each agent adopts the good they see in the action. This constitutes an intentional relativism, and according to Hegel, results in an unacceptable moral subjectivism.<sup>9</sup> Hegel's discussion of heteronomy of choice in *Natural Law* is more directly related to his emptiness charge against Kant. Hegel argues that, apart from choosing an end, the agent introduces a normative specification in terms of which his action is understood. Hegel contends, first, that this content is heteronomous and contingent. He then argues that it is this contingent content which undergirds the generation of a practical contradiction. In effect, the universalization procedure only works if the normative specification in terms of which the agent makes his action intelligible already presupposes the particular duty that the universalization procedure, in Kant's view, is designed to generate. I turn first to Hegel's moral subjectivity objection to Kant.

---

<sup>8</sup> "Whoever wills the end, so far as reason has decisive influence on his action, wills also the indispensably necessary means to it that lie in his power" (G, 4:417). For a discussion of Kant's containment relation in the context of the universalization test, see Korsgaard (1996, 94).

<sup>9</sup> This intentional relativism figures centrally in subsequent chapters, particular in regards to the Formula of Universal Law and natural actions (see chapter 2, 52ff) as well as the Formula of Humanity and positive duties (see chapter 5).

### 1. *Heteronomous ends and Moral subjectivity*

In §134 of the *Philosophy of Right* Hegel states that an “action for itself requires a particular content and determinate end.” This requirement is a consequence of Hegel’s more general view of action as the expression of a moral will (see *PR*, §113). For my purposes, the crucial feature of this view is that the moral will adopts a particular *intention* which grounds the *value* of the action. “The particular aspect of the action is its *inner* content ... this constitutes the *value* of the action and the reason why I consider it valid, i.e. its *intention*” (*PR*, §114).

This nature of action leads to a significant moral question; namely, *which* determinate end or intention should be ascribed to a given concrete action? This uncertainty concerning the intention of an action amounts to an uncertainty concerning which action an agent has willed. This uncertainty is disconcerting since, according to Hegel, “Even the *lowest degree* of understanding is enough to discover, like those learned theologians, a positive aspect in every action and hence a good reason and intention underlying it” (*PR*, §140). This worry is relevant because the Kantian may agree that the maxim ascribed to an action includes the intention of the agent. If an agent can discover a good intention to their action, and if they incorporate this good intention into their maxim, they will assess their maxim as good in light of this intention. This is problematic if *prima facie* impermissible actions may be judged permissible in virtue of the good intention the agent perceives in the action. For Hegel, to avoid an unacceptable moral subjectivity, there must be some restriction to the adoption of determinate ends or intentions and the ascription of maxims to action.

Hegel argues that Kant’s view cannot rationally restrict the adoption of intentions

and hence maxims. The autonomous rational will (*Wille*), as distinct from the arbitrary will (*Willkühr*), can only provide universal formal content. Thus, if we are to remain within the Kantian framework of Morality, and appeal only to internal content which the agent wills, then, according to Hegel, this leaves only heteronomous content furnished by the arbitrary will (*Willkühr*) of the agent. It follows that the intentional content of an action, an essential element of an agent's maxim, is an open question that can only be answered by the agent himself:

Since the action, a *determinate* volition, has a content, whereas *good in the abstract* determines nothing, it remains the task of particular subjectivity to give this abstraction its determination and fulfillment....[and] every subject is immediately accorded the honour [*Würde*] of providing the abstract good with content, or ... of subsuming a content under a universal. (*PR*, §140)

For any surd action there are various and conflicting ends, and the *moral character* of the action depends on which of these conflicting ends is subsumed: “But the subjective determination which I give to the good is the good I *know* in the action, i.e., my *good intention*. Thus, there arises a conflict of determinations, for one of them suggests that an action is good, whereas others suggest that it is criminal” (*PR*, §140). The agent himself determines what *action* is performed insofar as it is he who determines the end which his concrete action realizes. Consequently, the permissibility of an action is determined by the good – actually present or not – which each individual agent perceives in his action. Taken to its limit, this can justify any action whatsoever (see §140d).

To make this objection against Kantian Morality more concrete consider again the deposit scenario. An agent considers whether he can take or keep a deposit for which there is no proof. Kant ascribes to the agent the maxim to “to increase my wealth by every safe means” (*CPrR*, 5:27). In the remainder of the passage Kant makes clear that he

has in mind a maxim which satisfies and maximizes the individual's desire for happiness. Moreover, he thinks this purpose is destroyed when this maxim, as well as other maxims in service of individual happiness, is considered as universal law. (*CPrR*, 5:27-28)

We can read Hegel at this point as conceding this assessment of the agent's maxim, but with the caveat that this is a limited and conditional achievement. That the agent wills this end of increased fortune is entirely a matter of his own arbitrary will. The end which the agent adopts is entirely contingent, and it is entirely possible and even likely that this or another agent in similar circumstances would will a different end concordant with the good he perceives in the situation, and that, with this determinate end, his maxim could be universalized without self-contradiction.

Here is an example: An agent considers whether or not they can keep a deposit for which there is no proof, and in this concrete action operates on a maxim to 'use and distribute wealth and material goods in the most efficient and effective way available.'<sup>10</sup> Supposes that within the context of the deposit for which there is no proof, the agent sees keeping the deposit for himself and his benefit as realizing this end. He sees himself as achieving a good end. (To give him the benefit of the doubt, we may suppose that this is a maxim under which his overt action can really fall, rather than a false rationalization hiding his true but selfish maxim.) Moreover, it does not appear that willing to universalize this maxim with this end would lead to a frustration of his will.

---

<sup>10</sup> This is perhaps a grandiose way of expressing mundane intentions. One often hears excuses for actions such as, "They don't need this as much as we do" or "They're not really using this" or "With all the money they make, so and so does not really need this." As insipid as these excuses are in real life, they pose a real problem for the Kantian: these agents ostensibly see a valid end, the appropriate distribution of means and goods, in their wrongful actions. If we are to judge them by their self-ascribed maxim, as the universalization procedure asks us to, then we are forced to not only tolerate their actions, but to judge them permissible.

There is, on Hegel's view, no objective ground provided by Kant from which one may arbitrate between competing determinate ends. Rather, “it is the subjectivity of consciousness itself whose decision alone constitutes the objective element” and “[e]ven the *lowest degree* of understanding is enough to discover ... a positive aspect in every action and hence a good reason and intention underlying it” (*PR*, §140d). In this way “anything positive” about the concrete actions or the consequences of concrete action may be considered the 'essence' of the action, and that essence forms the good which may be adopted as the determinate end. But if any concrete action may be ascribed a good end, then any concrete action can be permissible. In this way, Hegel concludes that “any content one pleases can be subsumed under the good.” That is, any concrete action, no matter how apparently atrocious, may be judged “good” so long as the agent perpetrating the action is able to find in his action a “good” determinate end. The permissibility or impermissibility of a concrete action, then, is not determined by the universalizability principle, it is not determined by the autonomous will of the agent, but rather by the arbitrary will of each subject. For Hegel this constitutes an unacceptable moral subjectivism.

Is Hegel's moral subjectivity objection one that should worry the Kantian? After all, Hegel has fundamentally changed the focus of moral consideration from the agent's maxim to his concrete actions. In other words, the Kantian is concerned with whether or not a given maxim, with a particular end, can be universalized.<sup>11</sup> Hegel has changed the

---

<sup>11</sup> On this point Wood (1996, 156) presents a good reason as to why Hegel's objection is not one that should worry the Kantian:

Kant's four arguments from the FUL, if successful, show that one may not follow the specific maxims concerning suicide, false promising, and so on, which he mentions. They cannot show that it is always contrary to duty to commit suicide or make false promises, since those acts might be also done from quite different maxims which may pass the FUL test.

focus of moral judgment, such that our concern is now the ground, or lack thereof, for judging concrete actions in such a way that our judgment is not relative to the arbitrary ends adopted by each individual agent. However, this departure from Kant is not the product of confusion or a misunderstanding of the Kantian project. Rather, what Hegel seeks to demonstrate is that this transition beyond the testing and judgment of maxims by universalization is necessary, since an exclusive focus on the agent's maxim will have the counter-intuitive result of allowing any action whatsoever to be counted as good so long as the agent ascribes to his action a universalizable determinate end that he takes to be good. In other words, Hegel is forwarding an *external* objection to the universalization procedure. The universalization procedure is insufficient as a general approach for deriving positive moral duty insofar as it gives no satisfactory account as to why diverse acts generally agreed to be immoral are impermissible or contrary to duty. The procedure is only sufficient for determining our duty for a limited subset of possible intentions it itself does not specify, and yet permissible intentions may be attached by the agent to any number, and any kind, of egregiously impermissible *actions*.

## 2. *Heteronomy of choice and the superfluity of the universalizability test*

The focus of Hegel's discussion of heteronomy of choice in *Natural Law* parallels his arguments for the unacceptable intentional subjectivity of Kantian Morality. This

---

This interpretation of the Formula of Universal Law (FUL) presents an important rejoinder, and one which I cannot fully address here. I will only emphasize that Hegel's objection does not merely consist in pointing out that there are alternative, legitimate maxims; rather, Hegel's concern is the ground for adjudicating between competing maxims, and the worry that, for any context or situation, a subject may self-ascribe a maxim, based on his own perception of what is good, which justifies his action no matter how atrocious or vicious we may judge it to be. The main point is not that there are alternative legitimate maxims for various situations, but that there are *illegitimate* maxims for any situation.

discussion differs in that in *Natural Law* Hegel is primarily concerned with the normative specification by which we make our action intelligible. Unlike his moral subjectivity objection, this content does not consist in the end an agent adopts in his action, but in the normative concepts which are used and incorporated into the agent's will.

Understanding what Hegel means by “heteronomy of choice” in *Natural Law* and why this makes the imported subjective content illicit requires a careful reconsideration of the passage introduced in Section I:

But where is the contradiction if there were no deposits? The non-existence of deposits would contradict other specific things, just as the possibility of deposits fits together with other necessary specific things and thereby will itself be necessary. But other ends and material grounds are not to be invoked; it is the adopting either one specific matter or the other. For the form, however, one of the opposed specifics is just as valid as the other; each can be conceived as a quality, and this conception can be expressed as law. (*NL*, 77)

Hegel asks how a contradiction gets generated by the simple absence, or annihilation, of deposits, and answers that a contradiction can be generated *only if* “specific things” which constitute additional and imported “ends or material grounds” are “invoked” and “adopted.” Moreover, the material grounds invoked are contingent: different and opposing material grounds are equally valid from the standpoint of the “form” – that is, pure practical reason.

The controversial aspect of invoking material grounds can be seen in Hegel's contention that “one of the opposed specifics is just as valid as the other; each can be conceived as a quality, and this conception can be expressed as law.” By “quality” Hegel means that one of the “opposed specifics” is being conceived as *necessary*. Hegel is arguing that whether or not a contradiction is generated is conditional, and the condition that must be met is that one ‘specific thing’ be conceived as necessary and thereby



selected over competing and equally valid alternatives. In other words, invoking material grounds is not itself controversial; rather, it is implicitly invoking *these* material grounds and *conceiving* of them as necessary *qualities* that invites critical attention.<sup>12</sup>

To put this more concretely, to get the result that Kant thinks we must in the deposit scenario, the set of normative concepts which constitute the institution of property must “be elevated to something inherently necessary.” But this is a false elevation, and as such, Hegel contends that when the agent wills his action in terms of this necessary meaning of “property” he transforms conceptually “the conditioned and unreal into something unconditioned and absolute” (*NL*, 78).

Invoking as conceptually necessary what is contingent precludes any opposing conceptualizations from being invoked and thus used in the universalization procedure.<sup>13</sup> Since this privileging of one contingent conceptualization over another, and incorporating this conceptualization into the agent's will, is not grounded in pure practical reason, it constitutes a “heteronomy of choice.” Any determination of a maxim as universalizable or non-universalizable is thus conditional, and this condition is satisfied by the agent's arbitrary will. Prior to the universalization procedure one normative specification must be assumed and introduced into the agent's maxim and will. The only means of such an introduction is *Willkühr*:

Why must we attribute necessity to a particular normative specification? This is

---

<sup>12</sup> We find further support for this interpretation in the *Philosophy of Right*:

But if it is already established and presupposed that property and human life should exist and be respected, then it is a contradiction to commit theft or murder; a contradiction must be a contradiction with something, that is, with a content which is already fundamentally present as an established principle. Only to a principle of this kind does an action stand in relation of agreement or contradiction. (*PR*, §135)

<sup>13</sup> “If the one [set of material grounds] is fixed as absolutely subsistent, then...the other cannot be posited” (*NL*, 77).

required *if* we want to settle the question of permissible maxims via the universalization procedure. We can leave normative specifications contingent – and perhaps this is actually how we think of them – but we cannot think of them this way and still consider the universalizability test as sufficient for settling the pressing question of whether a proposed maxim is in accordance with duty. Said differently, leaving normative specifications conditional leads to an indeterminacy of duty. In order to resolve such indeterminacy, the normative elements (e.g., “property” and “deposit”) must be conceived according to a particular specification. The normative significance of these terms must be determined prior to the assessment of the agent’s maxim via the universalization procedure.

To see why this is we can reimagine the original scenario in the following way. Instead of an individual agent considering whether or not his maxim is permissible, consider two agents discussing whether one agent’s maxim is permissible. The second agent confronts the first and challenges the permissibility of taking the deposit. He makes his case by appealing to the universalizability test. The second agent, a good Kantian, accepts this test as the standard by which his maxim should be tested, but nonetheless insists that his maxim *is* universalizable. Eventually the agents realize that their dispute turns on their different understandings of the relevant normative concepts in terms of which the agent’s maxim is made intelligible. Now that the grounds of the dispute have been settled, how can the dispute itself be resolved? The agents would have to come to an agreement concerning the normative concepts; they would have to settle the meaning of “property” and “deposit” and the moral claims they make upon us.<sup>14</sup>

---

<sup>14</sup> We can put these points colloquially by considering the permissibility of deceit within exceptional circumstances, such as games like poker. Clearly a game is an artificial context where some rules of society

Hegel's point is that, in the original scenario Kant asks us to consider, this 'agreement' is done implicitly; one normative specification is being assumed as the *only* normative specification of "property." In other words, if everyone agrees on the norms governing the action, we can apply the universalization test and thereby confirm what is right or wrong. However, if there is any disagreement on what the governing norms are, then the test cannot settle this disagreement. It follows that the permissibility or impermissibility of a maxim is indexed to particular conventions – or to the conventional understanding of normative concepts adopted by the agent – but this leaves entirely open the question of duty for those agents who reject the norms in question.

One might object by saying that the normative specification incorporated into the agent's maxim is built into the situation, and so the use of this normative specification, as

---

are intentionally suspended. And yet, it is a rule of society (or at least our society) that one may make exceptions in the case of games. In other words, games may be artificial, but their very artificiality is a real and significant moral and social phenomenon.

Skilled poker players recognize that bluffing (i.e., intentional deceit) is an acceptable part of the game. Players have the *normative expectation* that their opponents will bluff. As such, we do not accuse a player of cheating when they bluff. Clearly, if the rules of the game were changed such that bluffing was prohibited, we would no longer have these expectations, and those who bluff would be accused of cheating.

Now we can imagine an untutored visitor to our society being invited to play poker and considering seriously whether he can intentionally deceive his opponent. A good Kantian, he employs the universalization test: What if everyone deceived? What if it were a *law* of the game to deceive? If deception were law, then no one would believe his current deception. It would seem then, that intentional deception is impermissible.

But bluffing in poker is *not* impermissible. Why is this? It is because there is a built-in exception to our normative specification regarding deceit: bluffing is allowed. And this is reflected in the change in terminology – we do not seriously accuse a bluffer of *deceiving* because the deceit was perpetrated within a game. In this way, the *meaning* of the normative concepts has changed to fit the normative context, and we understand what the agent does in terms of this transformed meaning.

Tutored in the ways of poker, our visitor now asks whether he can intentionally deceive his opponents. However, in considering his maxim and employing the universalization test he adopts normative concepts appropriate to his context: That is, in considering whether he can deceive his opponents, the agent now conceives of this deception in terms of the norms of the game and expectations of the other players; in short, he considers whether he can *bluff* his opponents. When universalized his maxim is seen as permissible. But this is a vacuous achievement insofar as the universalization test has only revealed that the agent may do, and only do, precisely what the norms of the game and society permit.

opposed to a competing specification, is justified.<sup>15</sup> The objection seems to concede to Hegel that what settles the permissibility of a maxim is the normative context, not the universalization test of the maxim itself. Furthermore, without an understanding, or implicit assumption, concerning the norms that are in place within a given context, we cannot come to any agreed conclusion concerning the worthiness of one's maxim. In other words, Kant holds that the "most common understanding...without instruction" (*CPrR* 5:27) can discern whether their maxim is permissible by appeal to the universalizability test, but Hegel shows this to be false. The agent must be well instructed; he must possess a practical facility with the norms in his situation in order to successfully employ the universalization procedure.

One might object that Hegel's critique, thus far, has failed to show that contingent normative content is illicit within the context of the universalizability test. As Wood has rightly pointed out and as Hegel agrees, no other material grounds are invoked other than the content of the agent's maxim itself. If the universalizability principle is designed to function as a test, where is the illicitness in an agent adopting a specific, even contingent, content? I have argued that this leads to an indeterminacy of duty and a relativization of maxims. However, this is not the end of Hegel's critique. A further illicitness of this conditional content can be seen by demonstrating the role this content plays in the derivation of relevant duties. That is, Hegel's critique goes beyond the contingency of normative concepts and argues that while the contingent specification of content incorporated into the agent's will does constitute sufficient grounds for generating a contradiction of will, this is only because the conclusion of the universalization procedure

---

<sup>15</sup> My thanks to Allen Wood for raising this objection in his comments to an earlier draft of this paper.

is *tautologous*. In this way, Hegel argues that a contradiction can be generated only if specific norms – which the universalizability test was designed to establish – are already at least implicitly assumed and incorporated into the agent's will. Thus, the crucial argument which supports Hegel's emptiness objection to the universalizability procedure is that the procedure is *superfluous*.

Returning to the *Natural Law* treatise, a consideration of additional passages provides both direct dialectical support for this crucial objection, as well as lending exegetical support to the interpretation I have proposed. Specifically, Hegel makes two arguments which contend, first, that the contradictions generated by the universalization procedure produce tautologies; and second, that *if* the particular “material grounds” imported into of the agent's will were changed, then the universalization test would produce different results. It is my view that we can only understand the contention that the universalizability test is a mere “empty formalism” in the context of these other arguments. The alternative interpretation here presented provides such contextualization.

The first of these subsequent arguments contends that if one normative specification of, for example, *property*, is assumed, then the universalization procedure will generate a contradiction, but the resultant conclusion will be a tautology:

If the specification of property in general be posited, then we can construct the tautological statement: property is property and nothing else. And this tautological production is the legislation of this practical reason; property, if property *is*, must be property. (*NL*, 77-78)

Hegel's analysis in this passage parallels his analysis of Kant's view of pure Understanding discussed above, in that he thinks pure practical reason can only produce tautologies from content provided independently. The passage also supports the view that Hegel's target is assumed conceptual content which is incorporated into the agent's will.

According to Hegel the generation of a contradiction will result if a “*specification of property...be posited.*” That is, if precise requirements of the concept of property are assumed, then a contradiction can be generated. Hegel's crucial point is that this contradiction is generated because the specification of property used already assumes the content that is to be derived via the universalization procedure; such a specification is illicit in that it leads only to a tautological legislation: “property, if property *is*, must be property.”

It what sense can “property, if property *is*, must be property” be understood as a tautology? This is not a modal tautology; that is, it does not follow from the mere fact that property exists that property *necessarily* exists. But from the context I do not think this is Hegel's meaning. Rather, Hegel means a normative tautology: If property *qua* normative convention is assumed (not to exist but to be true) then the norms of that convention hold; which is to say we are obligated to respect the norms of that convention. In the case in question, we must respect private property. This implies that any instance of property, for example a deposit, must be treated in accordance with the convention of property which has been assumed. In other words, if a normative specification of property obliges me to respect instances of property in a certain way, and I assume this normative specification of property, then I must respect instances of property as prescribed by that normative specification.

Simply stated, Hegel is saying that not keeping a deposit for which there is no proof is a norm, in this case a proscription, of the normative specification of property which has been posited and incorporated into the agent's will. Instances of property must be treated as property because the convention of property prescribes as much. But this,

according to Hegel, says nothing more than if one assumes a normative specification is true, and a normative specification consists in a set of norms, then one also assumes that these specific norms are also true.

To put Hegel's contention in more concrete terms, the concept of property *qua* moral institution incorporated into the will of the agent includes, as necessary attributes, specific principles and norms with regards to keeping and reporting deposits. Likewise, the concept of deposit incorporated into the agent's maxim consists in a type of property *qua* possession governed by the specific and contingent norms of the previously specified convention of property. In other words, property as a concept includes, among many things, a prohibition against keeping deposits for which there is no proof. A deposit just *is* property that by the current institution is entrusted for safekeeping, and which cannot be morally kept by another even in the absence of proof.

Applied to the universalization procedure, if we posit that property *qua* convention necessarily includes norms against keeping deposits for which there is no proof; and if a deposit just is an instance of property *qua* possession governed by these norms; then, inevitably, the willing that everyone *deny* having received a "deposit for which there is no proof" will generate a contradiction. But this is only because such a willing constitutes an internal contradiction with the specification of "deposit" incorporated into the agent's will. As such, the contradiction and conclusion are vacuous. What has been demonstrated is that if a particular though contingent concept of property which already includes a particular duty as a constituent part of the specification is willed, then willing a state of affairs contrary to this normative specification results in a contradiction. But this says only that an agent cannot consistently adopt the proscription

“One cannot appropriate a deposit for which there is no proof” while simultaneously willing the contradictory of this proscription “One *can* appropriate a deposit for which there is no proof.” Which is to say that an agent cannot will that both property *and* non-property obtain.

One might object that the concepts of property or deposit do not necessarily include the proscription that one may not keep a deposit for which there is not proof. Hegel agrees with this objection: the concept of property need not include this proscription. Hegel’s view is that the “specific things” or conceptual elements which constitute the normative specification incorporated into the agent’s will are contingent. However, the concept of property must be *conceived as* containing this proscription within the context of the action if the universalization procedure is to function successfully in this case.

Hegel provides a further, counterfactual argument for this contention: “[I]f we posit the opposite thing, negation of property, then the legislation of the same practical reason produces the tautology: non-property is non-property. If property is not to be, then whatever claims to be property must be canceled” (*NL*, 78). On a first pass Hegel appears to have in mind a normative specification of property without *private* property; for example, a normative specification of property where everything is held in common. If such a normative specification were assumed, we can imagine an agent considering whether or not they can realize their maxim to increase their fortune by any safe means through taking and keeping as private property what is held in common. If the content of the agent’s maxim in this context were universalized, we would see immediately that the agent’s end would be frustrated since there would be nothing held in common for the



agent to appropriate as private property. As such, an alternative normative specification reveals the “necessity” of a convention which contradicts the convention of private property. This reading is supported by Hegel’s further contention that “the aim is precisely to prove that property must be; the sole thing at issue is what lies outside the capacity of this practical legislation of pure reason, namely to decide which of the opposed specific things must be lawful” (78).

To this latter contention a Kantian may rightly object that proving “that property must be” is *not* the aim of the universalizability test. Rather, the aim of the test is to show that an agent may not exploit the norms of a given convention for his own benefit. The test shows the impermissibility of maxims which are exceptional in character regardless of the particular convention. In this way Hegel seems to have fundamentally misunderstood not the function, but the very aim of the universalizability test.

The Kantian objector is correct in characterizing the aim of the test. Hegel’s latter claim concerning the aim of the universalizability test appears to be an instance of reading Kant through the lens of his own theory. However, we should not allow this to detract from Hegel’s general critical strategy. Given what Hegel has argued in the preceding text, we can modify Hegel’s claim to mean that the sole thing which must *now* be at issue, given that the test produces only tautologies, is which normative specification must be lawful. In other words, if the universalizability test cannot be used to derive duties, but only to reveal those duties integral to an assumed normative specification, then the crucial ethical question becomes: Which normative specification ought we to adopt as lawful? Such a determination, according to Hegel, “lies outside the capacity of this practical legislation of pure reason.”

To see how Hegel's counterfactual argument is relevant to the aim of the universalizability test properly construed, let us consider an alternative convention that is not the opposite of property, but is nonetheless the negation of our normatively thick conception of property. Hegel holds that for the universalization procedure to function successfully, the agent must adopt as conceptually necessary a particular normative specification. As such, any alternative convention which does not include conceptual elements essential to this normative specification will constitute a negation of property so conceived. To illustrate Hegel's contention that the "negation of property" produces the tautology "non-property is non-property" let us incorporate such a specific negation of property into the agent's will and run the universalization test. Call this hypothetical concept of non-property 'Property-N'.

The alterations made to Property-N are deliberately *ad hoc*: Property-N is conceptually identical to what we call "property" except that it includes the established principle that any "deposit for which there is no proof" is nullified *qua* property, that is, it is non-property; furthermore, anyone may keep or deny having received non-property. Within Property-N the agent's original maxim becomes obviously permissible. The agent *may* deny having received a "deposit for which there is no proof" because a "deposit for which there is no proof" just *is* an instance of non-property, and such non-property is exempt from any property norms according to Property-N.<sup>16</sup>

This conclusion of permissibility, however, is obviously trivial. The conclusion states only that something which is nullified *qua* property, and is thus non-property, may

---

<sup>16</sup> One might object that Property-N is unintelligible. However, the alterations are not radically removed from our own conception of property, where documentation is required (whether, e.g., in retail, or real estate, or in a court of law) to establish the legitimacy of a claim. Property-N simply takes this requirement to its limit, such that there is *no* claim to something as property without documentation.

be treated as non-property. Which is to say that non-property may be treated as non-property; that is, kept in one's possession. This tautology results because the conclusion with its specific property norm and the specific norms of Property-N are identical; and the norms of Property-N as well as the concept of “deposit for which there is no proof” have been incorporated into the agent's will. It is this triviality that Hegel seeks to demonstrate. The apparent triviality is non-trivial or significant in that it shows how we do not need the universalizability test in order to discern this obvious conclusion; all we need is to understand the norms of Property-N. The point of this exercise is that the consequence of the original scenario – that the agent *cannot* keep the deposit – is equally trivial. It is trivial because this normative conclusion is a consequence of the normative features (“specific things”) of the convention posited and incorporated into the agent's will. Hegel is not arguing that the universalization procedure cannot not *function*; he is arguing, rather, that the successful functioning of the procedure is *superfluous*. “But pure reason demands that this shall have been done beforehand, and that one of the opposed specific things shall be presupposed, and only then can pure reason perform its now superfluous legislating.” (NL, 78; see also PR, §135 *Addition*)

Generalizing, this superfluity implies that determining which maxims *per se* are permissible or impermissible is not a matter of universalization. Any maxim may be determined to be permissible or impermissible *if* it is indexed to and construed in terms of the *right* convention or ethical practice. We can now see why Hegel would take the aim of ethical deliberation to consist in deciding “which of the opposed specific things must be lawful.” If the permissibility or impermissibility of each given maxim is indexed to a given ethical practice, then the important normative work to be done is to determine

which ethical practice we ought to adopt *such that* proposed maxims may be determined permissible or impermissible. In other words, the important question is which ethical life we should adopt, the terms of which will determine how we understand ethical maxims and concrete actions.

### III. OBJECTIONS

The above alternative interpretation faces two principle objections: First, that the results are trivial; and, second, that the Formula of Universal Law and universalization test do not concern particular kinds of actions, but rather a particular character of maxims; namely, maxims which are exceptional with reference to the convention in which they operate.

#### A. Triviality Objection

The above explication already included one sense of triviality to the emptiness objection newly interpreted: that the maxim to “keep a deposit for which there is no proof” becomes permissible when the proscription against keeping deposits is absent from the counterfactual normative specification is trivial in that the permissibility is a direct consequence of permitting the maxim. I argued, however, that such triviality is illustrative of Hegel’s position and equally attributable to the original case and normative specification.

Another form of triviality concerns not the grounds of permissibility, but rather the fact that in the counterfactual normative contexts the maxim is permissible and, intuitively, ought to be permissible. That is, the permissibility of keeping a deposit for which there is no proof within Property-N is not a false-negative. One might infer from

this that the universalization procedure functions as it should and while it may permit opposing actions in counterfactual convention, it does not permit immoral actions. In this way, Hegel's objection is trivial in that it only shows that counterfactually permissible maxims are possible. Simply, an example has not been furnished where the universalization procedure fails to generate a contradiction within a counterfactual normative contradiction and, thereby, fails to proscribe an immoral maxim.

This shortcoming is readily remedied. A relatively minor example is that of queuing-up. If I am in line at the train station and I consider the maxim "in order to get a better seat, I will push to the front" one can easily see that the maxim would generate a contradiction: if everyone in line were to do so my maxim would (most likely) be frustrated. This proscription is, moreover, antecedently posited in the normative specification of queuing. Importantly, the maxim is plausibly immoral on the grounds that it is inconsiderate, not merely in the sense of violating etiquette, but in the most literal sense of not considering the interests of others and putting one's own interests above the interests of others. Everyone has an equal interest in finding a good seat, but I give no heed to this equality of interest. Finally, we can easily see that the maxim is permissible, that is it does not generate a contradiction, if we change the normative specification such that queuing is not part of the normative convention at train stations. Crucially, however, the impermissibility remains: in the absence of the relevant norm I propose that the action of pushing to the front is still inconsiderate (one elevates one's own interests above that of others who have the same interest). Though this is a minor example, it is of particular import because it also challenges the determination of impermissible maxims in virtue of being exceptional in character if this is understood as exploiting a convention for one's

own advantage. In the example of pushing to the front of a line, the agent does not make an exception of himself (everyone may and does do the same), but merely does not concern himself with the interests of others.

More profound examples are available as well. Freyenhagan (2012) provides an example of buying slaves to free them (the maxim of which would be frustrated if universalized) that can be easily modified for our purposes. Suppose a businessman finding himself losing to his competitors proposes the maxim “in order to gain an advantage over my competitors, I will force workers to work without pay and keep them in bondage.” Such a maxim certainly would be expected to generate a contradiction in conception; however, the success of the universalization test in this case is complicated. If forcing another to work and live in bondage is conceived as a natural action, then no contradiction can be generated through undermining a convention. However, there are also contingent conventional grounds for generating a contradiction in this case: the businessman’s motivation is to escape a current financial disadvantage (similar to the false-promising agent who seeks to escape a financial difficulty). The relative financial advantage of slavery depends on the conventional practice of wages: if everyone in the same circumstance were to follow the maxim,<sup>17</sup> that is if wages were to be eliminated, then there would be no distinct competitive advantage to enslaving his workers and the businessman’s specific end would be frustrated. Thus, the maxim is impermissible in virtue of the normative specification in which it operates. That is, the proscription produced from the universalization test is superfluous.

This superfluity can also be seen by positing counterfactual normative

---

<sup>17</sup> We must assume, of course, the appropriate empirical and, in this case, socio-economic conditions obtain, which is true for universalizing every maxim of a conventional action.

specification where slavery is permitted. Imagine a second businessman who pays his workers (out of benevolence or savvy), but finds himself facing financial difficulty and so proposes to himself the same maxim of the first businessman: “in order to gain an advantage over my competitors, I will force workers to work without pay and keep them in bondage.” In this second case the motive is the same as the first businessman (“to gain an advantage”) as well as the specific means (“force workers to work without pay and keep them in bondage”); however, the reasoning must be different. The first business man reasons that he will gain an advantage because his position will be *exceptional*. The second businessman cannot reason this way for the simple fact that by enslaving workers he will not gain an *unfair* advantage over others who are required to pay wages; slavery is permitted and others may and do choose to adopt the same maxim when and if they see it as to their advantage. The second businessman, rather, reasons that slavery is the financially judicious option given his current challenges and needs.

In other words, the first business man adopts slavery as a means for exceptional reasons, and the exceptionality of the reasons make the maxim non-universalizable. The second business man adopts slavery as a means for advantageous but not exceptional reasons. Why slavery is an exceptional rather than merely advantageous means, however, is a contingent matter of convention and antecedently determined normative specifications. Simply, the first businessman's maxim is non-universalizable precisely for the reasons it should be: it is exceptional in character. However, if we change the character of the convention, the reasons for a maxim change with it and the same, grossly impermissible maxim may be permissible. Of great interest is the fact that a profoundly impermissible conventional maxim is not impermissible because it is exceptional in

character.

The question of whether such an action conceived in natural terms is permissible is, for now, still open – I address this further question in chapter 2. However, construed as a conventional action, the contradiction generated through the undermining of the convention of wage-earning is not generated in a counter-factual convention where slavery is practiced. Within such a convention a maxim of slavery is evidently permissible: one can force one's current workers to work without pay and within bondage.<sup>18</sup> One cannot employ the maxim for precisely the same reason (*viz.*, to gain an *unfair* advantage over one's competitor); however, under the right contexts one could adopt the maxim with, for example, the goal of maximizing one's advantage. Crucially, a maxim of slavery, which is profoundly immoral, is universalizable.

## B. Hypothetical Universality Objection

There is a second important objection to the foregoing position. As noted above, on some interpretations of Kant the *aim* of the universalizability test is not to “to prove that property must be” or that any other normative practices, such as promising, ought to be adopted. Rather, the aim of the test is to show that, for *any* normative practice or convention, regardless of how it is specified, the exploitation of that practice for personal

---

<sup>18</sup> In chapter 2 I consider O'Neill's contention that both the slavery and slave-holding maxims are non-universalizable and impermissible. As is the case throughout the dissertation, much depends on the description of the maxim. In the case of slavery and slave-holding, the agents are proposing to become slaves or slave-holders, and thus, are presented as not currently either. In this case, universalizing the maxims is evidently impermissible since it is equally impossible for everyone to abrogate their autonomy or for everyone to be masters. In the current case, however, the agent is considering a maxim of slave-holding that is subsequent to being a businessman or slave-holder. The situational specificity is essential to the formulation of the relevant maxim and, consequently, the permissibility of the action.



or individual gain is impermissible. In other words, maxims indexed to any normative practice which are “exceptional character” are impermissible.<sup>19</sup> Another way of stating this is the agent in the deposit scenario does not simply will a world with property, and a world without property; he wills that he possess the deposit he intends to keep *as* property governed by the same norms he simultaneously wills *not to exist*. As such, all the test needs to “prove” is that the maxim the agent wills is exceptional in character relative to the convention and norms of property that the agent himself wills to uphold.

Said differently, the particular examples considered by Kant only serve to exemplify the more fundamentally impermissible attributes of some maxims. In the case of the deposit, what the agent fundamentally exploits is the *trust* of the other agent who made the deposit. Under any alternative normative specification that does not include trust concerning deposits without proof, the agent would not be exploiting trust since there would be no instance of entrusting. Accordingly, by this use the universalizability test would be neither vacuous nor superfluous, in that it would serve to determine whether or not a maxim indexed to a particular ethical practice violates and exploits a more fundamental ethical relation of trust regardless of the particular convention in terms of which the agent is acting. In this way, the universalizability test establishes *universal* requirements on all maxims regardless of the prevailing ethical institutions. And the particular proscription against, say, keeping a deposit for which there is no proof is only

---

<sup>19</sup> See Korsgaard (1996, 15) who apparently has Hegel in mind when she writes:

Later critics claim that undermining the efficacy of promises is only a contradiction if promises are themselves necessary. But Kant's point in the example is more modest than that; it is not intended to establish that promises are necessary. Promises are necessary for the man in the example, because *he* proposes to use a promise as the means to his own end. This is why Kant says that he cannot will his maxim and its universalization *at the same time*. Whenever you propose to perform an action whose efficacy depends on its exceptional character, you get a contradiction of this kind.

an instance of these universal requirements of all ethical life.

I will refer to this position as the hypothetical universality interpretation.<sup>20</sup> It states that when an agent wills to participate in any given ethical practice or have the norms of any given practice apply to them, then a maxim which is found to be exceptional in character relative to that practice is impermissible for the agent. On this view, there is one thing an agent ought not to do regardless of his convention, and this is to make himself an exception by means of exploiting the norms of that convention.<sup>21</sup>

I will respond to and reject this objection first in terms of trust and, in doing so, will again focus exclusively on the deposit scenario. Trust, or the violation of trust, ungirds both the deposit and false-promising examples, and is a useful heuristic for illustrating maxims which are exceptional in character, and thus the hypothetical universality interpretation as well as the Hegelian response.<sup>22</sup>

According to the hypothetical universality interpretation any maxim, within any institution of ethical life requiring trust, can generate a practical contradiction in the

---

<sup>20</sup> My thanks to Peter R. Murray and Baruch Brody for bringing this interpretation to my attention.

<sup>21</sup> While this is an important and formidable objection, it is also worth noting how far we have already come, and how far removed this modified Kantian position is from Kant's own desideratum. As will be discussed in subsequent chapters, Kant not only sought to show the impermissibility of exceptional maxims relative to given conventional norms, he also sought normatively substantive and particular conclusions regarding certain action types. So, for example, in the *Metaphysics of Morals* Kant derives particular proscriptions against acts of suicide as well as acts of *untruthfulness*. This latter proscription is particularly illustrative of Kant's intention to prohibit particular kinds of actions: untruthfulness is wrong *per se*, and not because it represents actions or maxims that are exceptional in character. Admittedly, Kant's proscription against untruthfulness is within the context of ethical duties owed to oneself; one might, then, think that the criterion of exceptionality could more consistently apply to duties of right or justice. Perhaps, though Kant's apparent rigorism concerning maxims of deceit in the context of right militates against this view.

<sup>22</sup> Maxims that violate trust exemplify maxims which are exceptional in character; though importantly, not all maxims which are exceptional in character instantiate violations of trust – or at least it is not clear to this author that they do. For example, impermissible maxims of natural actions, such as murder or thievery which do not involve a convention such as deposit-making, also exemplify maxims which are exceptional in character but do not, *prima facie*, involve the violation of trust.

following sense: An agent cannot consistently will to violate the trust of others in any particular way, while simultaneously willing a condition or state of affairs which undermines the particular form of trust which others have given.<sup>23</sup> Doing so would exploit the trust which has been given.

Hegel need not deny this moral truth so long as it remains *abstract*. That is, Hegel can accept this Kantian rejoinder and still contend that it provides no concrete practical guidance until a normative content is posited which informs us of the particular way trust may not be violated. *Which is tantamount to telling us what trust is.* Furthermore, to adopt the arguments from the previous section, Hegel's point is that the particular way trust cannot be violated is determined heteronomously and prior to using the universalization procedure. The universalizability principle tells us that we may not simultaneously will to be trustworthy and untrustworthy; but it is only the adoption of a particular convention that tells us *if* and under what conditions we *are* untrustworthy.

Pithily, what the universalizability test shows is that *betrayal* is impermissible. However, without content which the test cannot itself provide, this is no more informative than saying that *murder* is impermissible. Of course *murder* and *betrayal* are wrong; but these are wrong *abstractly*, by definition. There is nothing in the abstract concept of betrayal, which just is the impermissible breaking of trust, which can specify which particular action constitutes or falls under the concept of betrayal. To achieve more than a truism, we must determine when and under what condition an agent has betrayed another's trust.

To see more concretely how Hegel's emptiness critique applies to fundamental

---

<sup>23</sup> My thanks to Allen Wood who presented this Kantian rejoinder in his commentary to an ancestral version of this chapter.

moral categories, we may consider a simple and highly abstract example of trust.

Consider two agents, Anna and Henry. Anna entrusts Henry with X. At this point, no particular convention or norms are assumed, and we have not yet specified what X is. We now ask, “Can Henry break Anna’s trust?” (Notice that I did not say “betray” Anna’s trust. “Betray” is a normatively laden term; it already implies wrongdoing and, thus, impermissibility. We need to remain neutral at this point to discover under what conditions the neutral “breaking” of another’s trust constitutes betrayal.)

We may determine whether Henry can break Anna’s trust by asking whether Henry’s maxim to break Anna’s trust can be universalized. The property and promising examples in Kant all give self-interest, or benefiting oneself, as the motive behind the agent’s maxim. This, as was seen in the above discussion of heteronomy of choice, need not be the case: We can ask after the permissibility or impermissibility of breaking another’s trust for any number of motives. So we might first ask whether Henry can break Anna’s trust for *her* benefit; and then ask whether Henry can break Anna’s trust for *his* benefit.

In both cases Hegel’s answer is that we do not and cannot yet know whether Henry can ethically break Anna’s trust. The Kantian may and should respond that we can know: If Henry’s maxim to break Anna’s trust were universalized – that is, if it were made a *universal law* that Henry would break Anna’s trust – then Anna would never have trusted Henry to begin with. In other words, if it were a law that Henry breaks Anna’s trust, then Anna would never trust Henry because there would simply be no trusting.

However, under what specific conditions is this true? Is this true under all conditions of trust? Is there not a conceivable practice, or agreement, of trust where Anna

trusts Henry with X, and would continue to trust Henry with X *even if* he were to break this trust for his or her benefit? To clarify this point that certain conditions must be met before we can know whether or not Henry can consistently will his maxim, some particulars need to be introduced.

Consider first that Henry's maxim is to break Anna's trust in order to benefit her in a situation where the benefit in question is saving Anna's life. Would Anna still trust Henry with X if Henry's maxim to benefit her greatly by saving her life were made the law? We can easily see how Anna *would* continue to trust Henry with X if such a maxim were law. Anna's trusting Henry with X would be understood to include the exception that if her life were in danger (or if there were some other great harm to her) Henry could break her trust.

Now consider a maxim where Henry benefits. Can Henry universalize a maxim to break Anna's trust in order to benefit him greatly, say in a situation where he would receive a large sum of money through breaking Anna's trust? The answer here is the same as above. We can easily conceive of a normative expectation shared by Henry and Anna such that there is an exception to her entrusting Henry with X such that if he can benefit himself greatly then he may break Anna's trust.

The Kantian here could rightly object that in these scenarios Henry is not breaking Anna's trust at all, insofar as he is abiding by the *conditions* of Anna's trust, by, we might say, the conditions of the practice or convention of trust Anna and Henry share. This is precisely right, and this is what constitutes the distinction between "breaking" and "betraying." In these cases we can say that Anna's trust was *broken* but it was not *betrayed*.

This distinction illustrates precisely Hegel's point: The impermissibility or permissibility of Henry's maxim – whether or not it is *exceptional in character* – precisely maps on to the conditions that obtain in the practice of trusting which Anna and Henry share. Crucially, these conditions are not grounded in Henry's pure practical reason; Henry implicitly *incorporates* these conditions into his will and maxim, but the specification of the conditions constitutes external material grounds of heteronomous origin. In simple terms, we can only know whether or not Henry's maxim is exceptional in character, that is, whether or not he has betrayed Anna's trust, *after* we ask what Anna's normative expectations are (or what Henry understands Anna's normative expectations to be). Moreover, with this requirement in place, the process of universalizing Henry's maxim in order to determine whether or not it is exceptional in character, and thus impermissible, is superfluous confirmation of what he is presumed already to know, or what could be known by simply asking what the normative expectations are. Since the universalizability or non-universalizability of Henry's maxim maps onto to the conditions or practice in place, all that is needed in order to know whether or not Henry can permissibly follow his maxim is to ask what the conditions of trust are – what are the shared normative expectations regarding the keeping of trust. Once this is determined, what is permissible for Henry, what duty requires and allows with regard to trust, is precisely “what is prescribed, expressly stated, and known to him in his situation” (*PR*, §150).

Thus, returning to the hypothetical universality objection, we can reframe Hegel's general argument as contending that whether or not a maxim is exceptional in character is determined by which exceptions are built in to the normative specification in terms of

which the concepts integral to the maxim are understood. But this is just to say that the exceptionality of a maxim is determined by content that is itself determined prior to and independently from the “practical legislation of pure reason.”

Which exceptions should we build into a normative specification? When is it permissible to break another's trust? Hegel views these apparently distinct questions as having the same answer. Whether a maxim is exceptional in character, and thus impermissible, depends on and is symmetrical with the exceptions built into the normative specification in terms of which a maxim itself is understood. When we recognize that whether a maxim is exceptional in character depends on whether the exception is built in to an antecedent convention, then we see that any particular maxim, given the right normative contextualization, can pass the universalization procedure. Hegel writes that “there is nothing whatever which cannot in this way be made into a moral law”<sup>24</sup> (*NL*, 77). The permissibility or impermissibility of any maxim removed

---

<sup>24</sup> Hegel makes a similarly extreme contention in the *Philosophy of Right*.

One may indeed bring in material *from outside* and thereby arrive at particular *duties*, but it is impossible to make the transition to the determination of particular duties from the above determination of duty as *absence of contradiction*, as *formal correspondence with itself*, which is no different from the specification of *abstract indeterminacy*; and even if such a particular content for action is taken into consideration, there is no criterion within that principle for deciding whether or not this content is a duty. On the contrary, it is possible to justify any wrong or immoral mode of action by this means.” (*PR*, §135)

Some view this extreme contention as clearly absurd. For example, Wood writes: “Yet Hegel is surely wrong to say that this is a test that any maxim can pass. There definitely are principles that one person can follow, but that it would be contradictory to suppose that all follow: For example, ‘I will never work, but always live by exploiting the labor of others.’ If everyone followed the principle, there would be no laborers left to labor for others” (Wood, 1990, 157). I take Wood to mean here that the universalization procedure shows the proposed maxim to be exceptional in character and, thus, impermissible. Moreover, I take Wood's point regarding Hegel to be that, regardless of what success the emptiness charge may have generally, Hegel is wrong in his extreme view that any maxim whatsoever could pass the universalization test.

Let us reconsider Wood's proposed maxim: “I will never work, but always live by exploiting the labor of others.” Having such strong Kantian intuitions myself, I initially found Wood's judgment to be correct that this maxim to be incontrovertibly impermissible. However, on closer examination this example is much more complex than it initially appears to be and a judgment of impermissibility is deeply mistaken. First, Wood has selected a natural (i.e., non-conventional action) as his example. This

from its normative context is indeterminate. What will determine the permissibility or impermissibility of the maxim *per se* is the normative specification in terms of which the maxim is understood. Moreover, once it is conceded that exceptions and conditions may be built into a normative specification, and that this process lies outside of the legislation of pure practical reason, then we recognize that a normative specification may include *any* exceptions and conditions whatsoever. Simply put, whether or not a maxim is

---

complicates things, as with natural actions no normative specification need be posited in order to perform the action, and thus a contradiction cannot be generated by undermining the necessary conventional means of performing the action. In this way, natural actions present a peculiar problem for the universalization test as well as the application of Hegel's emptiness critique. Indeed, Hegel's objection to this maxim would mirror his objection to maxims of positive duty (such as "Help the poor") and his contention that universalization test in such cases "shows too much" (see chapter 3). This maxim of excepting oneself from work also shows to much – that is, it is excessively proscriptive.

To see this, we must, second, clarify the terms of the maxim, as "work" can be interpreted in various ways. From the context and his pairing of "work" with "labor," I think Wood is construing "work" in terms of productive contributions within an economic system. If this is correct, then there are uncontroversial exceptions to the impermissibility of this maxim. Ordinary examples would include spouses and heirs: a spouse who, for example, stays home to raise a family or perhaps just takes care of domestic concerns may permissibly exploit the labor of others. Similarly, with the condition that the inheritance was not of ill got wealth, an heir in virtue of his inheritance may literally live off of the previous work of others. Additional examples of exceptions include adults for whom any work is excruciatingly painful, or who are severely disabled; also, minors with terminal diseases who will, with all probability, die before or early into adulthood. If we accept these exceptions, then the impermissibility of willing this maxim is restricted to a subset of agents. If we do not accept these exceptions, then the universalization test will produce false-positives. The implication is that while we may be able to find maxims that cannot pass the universalization test, this is not a sufficient condition for showing the permissibility or impermissibility of the maxim; furthermore, that a maxim is *exceptional in character* is also not a sufficient condition for determining the permissibility or impermissibility of a maxim – some forms of exception and exploitation are permissible.

If this is right, then Hegel's view presents the challenge of justifying this restriction: in order to determine our actual duties, we must bring in content from outside. In other words, if some exceptions or conditions may be built into a normative specification such that this maxim is permissible for some agents, and if the incorporation of these exceptions is not the legislation of pure practical reason, then there is no rational restriction on the incorporation of other exceptions and conditions into competing normative specifications. Moreover, the incorporation of alternative exceptions and conditions into competing normative specifications means that the proposed maxim may be "permissible" for these agents and in ways which we would reject.

Thus, for example, we might consider a normative specification where there is a work exception for artists. Any artist may live off the work of others while never laboring themselves. (Again, from the context I am assuming that the relevant sense of "work" here is one that pertains to the production of material goods, or that contributes to an economy in such a way that enables us to live.) Likewise, an exception may be made for the beautiful, the humorous, or athletic. Simply, once any exception is permitted, then any other exception may also be permitted in a competing specification. And what is permissible will depend on which exceptions have been antecedently and heteronomously incorporated.



impermissibly exceptional in character depends on which exceptions are permitted, and which exceptions are permitted is a question the universalizability test cannot answer.

#### IV. CONCLUSION

The objective of this chapter has been to provide a more tenable interpretation of Hegel's emptiness charge by showing how the emptiness objection must be understood in the context of supporting arguments. Hegel's emptiness charge, consistently dismissed as based on a confused interpretation of the universalizability principle, has been reinterpreted and shown to target intentional and normatively specific conceptual content already incorporated into the agent's will and maxim. Hegel contends that the specification of norms necessary for generating a contradiction via the universalization test are such that the conclusion of the test is tautological, and thereby trivial. Furthermore, this triviality renders the universalizability test itself superfluous.

The emptiness objection itself, that is, the contention that the universalizability test is a mere formalism, becomes an objection only after Hegel establishes that the particular institutional or conventional content is heteronomous and has a normative specificity which renders the universalization procedure superfluous. In other words, Hegel's emptiness objection, properly construed, is that the empty formalism of the universalizability test is what is *left* after what is for Kant illicit content has been removed from the universalization procedure:

These determinations [of particular content and a determinate end], however, are not contained in the determination of duty itself. But since both of them are conditional and limited, they give rise to the transition to the higher sphere of the *unconditional*, the sphere of duty. Hence all that is left for duty itself, in so far as

it is the essential or universal element in the moral self-consciousness as it is related within itself to itself alone, is abstract universality, whose determination is *identity without content* or the abstractly *positive*, i.e. the indeterminate. (PR, §135)

Without a content posited by the arbitrary will, in the form of both intentions and a normative specification, our particular duty is left indeterminate. That is, though we may know that we have a duty to others, that we ought to act from duty and promote the good (PR, §§133 and 134), and that our duty accords with abstract *positive* moral categories (such as *trust*, *benevolence*, *honesty*) and is contrary to abstract evil categories (*deceit*, *murder*, *betrayal*, *dishonesty*) (see PR, §140d); we are still powerless to realize particular duties (see PR, §141 *Addition*). “For the universal aspect of good, or good in the abstract, cannot be fulfilled as an abstraction; it must first acquire the further determination of particularity” (PR, §134, *Addition*).

Within this indeterminacy of our particular duty, our duty can only be understood formally: Our individual maxim must not be contradictory with this same maxim willed as universal law. Without appeal to particular content, this formal imperative consists in nothing beyond a practical principle of non-contradiction. When both the individual will and the universal will are considered, a maxim to *deceive* is impermissible because the agent simultaneously wills to be believed and not believed; a maxim of *betrayal* is impermissible because the agent simultaneously wills to be trusted and *not* to be trusted. And yet when and under what conditions one has impermissibly deceived or betrayed must be determined antecedently and independently from this formal imperative. Without particular content the universalization procedure remains abstract, and therefore empty.

In order to know our particular duty, we must introduce both intentional content and a normative specification in terms of which our action is made intelligible. In order to

know whether my deception is impermissible, the particular norms governing trust, promising and similar conventions must be determined. That is, the *meaning* of moral terms must be established before a substantive judgment of a moral action may be made.

We can see even from the narrow and limited focus of these arguments the broader trajectory of Hegel's project, and specifically why Hegel contends we must transition from *Morality* to *Ethical Life* – that is to *Sittlichkeit*. Crucially, *Sittlichkeit* is not the abandonment of Morality for ethics; but rather the fulfillment of Morality and actualization of duty through the incorporation of a particular ethical life: “In ethical life as a whole, both objective and subjective moments are present, but these are merely its forms. Its substance is the good, that is, the fulfillment of the objective [united] with subjectivity” (*PR*, §141 *Addition*). Within a *Sittlichkeit* we may now know when and under what conditions our maxim is permissible (see *PR*, §150). By unifying Morality with ethics, *Sittlichkeit* provides the content necessary for actualizing our duty (see *PR*, §149).

## **Chapter 2:**

### **Further Consideration of the Relevance of Hegel's Emptiness Critique for the Formula of Universal Law**

Chapter 1 introduced and developed an alternative interpretation of Hegel's "emptiness objection." This alternative interpretation maintains that the emptiness of the Formula of Universal Law is not an independent objection, but rather part of a constellation of arguments which form an emptiness *critique* of Kant's Categorical Imperative. Moreover, I argued that Hegel explicitly recognizes how the structure and function of the universalization procedure includes a purely formal element as well as heteronomous content from the agent's own maxim. In this way, I argued that Hegel's emptiness objection parallels his moral subjectivity objection in that both objections contend that the successful function of the Formula of Universal Law necessitates importing illicit content. For these reasons, I argued further that it would be contrary to Hegel's own explication of the universalization procedure to contend merely that the Formula of Universal Law is an empty formalism, and for this reason alone cannot be used to determine particular duties. Rather, as I argued, the crux of Hegel's critique targets the content of the agent's maxim and heteronomous will which is utilized in and essential to the universalization procedure. In this way, the formal emptiness of the universalization procedure only constitutes an objection when understood as a consequence of Hegel's overall critique.

With this new understanding of Hegel's critique in hand, chapter 1 established

that Kantian rebuttals have dismissed Hegel's charge of emptiness without understanding how the emptiness charge functions within a comprehensive critique. In short, Kantians have responded to a straw man, and Hegel's full critique remains both devastating and unanswered.

Given the new interpretation of Hegel's emptiness critique, and the failure of Kantian rebuttals to accurately identify or respond to this critique, the question remains what implications this alternate interpretation has for the Categorical Imperative and Kantian moral philosophy generally. The remainder of the dissertation is dedicated to answering this question. The current and following three chapters of the dissertation explicate the extent and consequences of Hegel's critique properly understood. Significantly, this contention is not limited to the Formula of Universal Law. Hegel's emptiness critique has been criticized<sup>1</sup> for focusing too narrowly on the Formula of Universal Law, and not addressing other formulations of the Categorical Imperative or how these other formulations may supplement the Formula of Universal Law. While it is true that Hegel does not make explicit how his critique bears on other formulations of the Categorical Imperative, the arguments which result in the emptiness of the Formula of Universal Law are readily extended to and equally effective at undermining these alternative formulations – or so I shall argue in chapters four and five.

Presently, however, I will continue to focus on the Formula of Universal Law. This is necessary for two reasons. First, the discussion of Hegel's critique in chapter 1 was restricted to a single canonical example: keeping a deposit for which there is no proof. The example of the deposit served as a useful heuristic for explicating Hegel's

---

<sup>1</sup> See, e.g., Lo (1981) and Freyenhagen (2012)

critique, but leaves in question whether and how Hegel's critique applies to other particular duties. In other words, it remains unclear whether Hegel's critique can be extended to other canonical examples, and crucially maxims of both negative duties and positive duties. Second, the Formula of Universal Law has various and contrary interpretations, and thus the full relevance of Hegel's critique, even when restricted to the Formula of Universal Law, has not been established.

Demonstrating the effectiveness of Hegel's critique for all interpretations of the Formula of Universal Law will inevitably be limited; debate continues regarding not only the kind of contradiction the Formula of Universal Law is intended to generate, but also the very nature and scope of maxims. Accordingly, I limit my discussion to two influential and representative interpretations of the Formula of Universal Law. Specifically, I first continue to consider the relevance of Hegel's critique for Christine Korsgaard's Practical Contradiction Interpretation. I then consider Onora O'Neill's interpretation of maxims as "governing principles."

This circumscription<sup>2</sup> is thrice justified: First and foremost, both Korsgaard and O'Neill consider their interpretations to be immune to Hegel's critique. Second, both Korsgaard's and O'Neill's respective positions, though developed in different contexts, are related to the Act-Description Problem which concerns how an action is described

---

<sup>2</sup>Barbara Herman and specifically her (1993) influential book *The Practice of Moral Judgment* as well as her (2008) collection *Moral Literacy* could also be included here. I have not included Herman's account for two key reasons: first, the relevance of Hegel's emptiness objection for Herman's position has been thoroughly addressed in two essays: Sedgewick's (2011) '*Letting the Phenomena In*': *On How Herman's Kantianism Does and Does Not Answer the Empty Formalism Critique* and Kenneth Westphal's (1995) *How "Full" is Kant's Categorical Imperative?*; second, tenability of Herman's position, including "rules of moral salience" and the functional relevance of a "middle theory" depend on the contentfulness of the Formula of Humanity and its ability to safeguard the Formula of Universal Law from charges of emptiness. I take up the emptiness of the Formula of Humanity in chapters 4 and 5.

and, consequently, the formulation of the relevant maxim. Korsgaard's Practical Contradiction Interpretation is particularly influential within contemporary Kantian ethics because of the important solution it provides for the problem of natural actions. Third, and related to the second reason, Korsgaard and O'Neill develop representative views of two opposing interpretations of how maxims ought to be construed. Korsgaard construes maxims so that they incorporate an agent's specific intention<sup>3</sup>; that is, maxims are construed *thickly*. In contrast, O'Neill maintains – for reasons very close to Hegel's objection – that maxims cannot tenably include the agent's specific intentions, but must be construed as broad “governing principles” – that is, maxims must be construed *thinly*. Considering these views, then, provides a useful contrast of different solutions to the Act-Description Problem as well as a contrast case for how Hegel's critique can be appropriated for varying interpretations of the Formula of Universal Law.

## I. THE PRACTICAL CONTRADICTION INTERPRETATION: FURTHER IMPLICATIONS OF HEGEL'S CRITIQUE

The Practical Contradiction Interpretation has already been discussed in chapter 1, and forms the foundation to the Kantian rebuttal to the standard interpretation of the emptiness objection. This rejection turned on the contention that the contradiction to be generated by the universalization procedure is an internal contradiction of will: the agent generates a contradiction by willing a world, say, in which he makes a promise which is

---

<sup>3</sup> Intentions can be formed at various degrees of specificity. I do not mean to imply that Korsgaard maintains that maxims must be construed at the highest degree of specificity; but rather that in formulating maxims Korsgaard incorporates intentions unique to the specific act in question (e.g., the intention to *secure a particular position*), and this thick construal of maxims, which is distinct from maxims *qua* governing principles, brings with it implications for how Hegel's emptiness critique may be applied.

respected, and simultaneously willing a world in which false promises are universally made, and consequently any earnest attempts at promising are ridiculed. What is unique to the Practical Contradiction Interpretation is that the contradiction generated is understood in terms of the agent willing two mutually incompatible means, and thus as a *practical* contradiction.

We have seen in chapter 1 that this interpretation of Kant does not avoid Hegel's critique. And yet, the extent to which Hegel's critique remains relevant has not been fully made clear. This is due, in part, to the goals of chapter 1, but also the broad relevance of the Practical Contradiction Interpretation. In addition to its strictly exegetical appeal, the Practical Contradiction Interpretation is noteworthy for providing a solution to the problem of natural actions. Accordingly, my goals in this section are two: First, I must show that Hegel's critique is generally applicable, and so will apply his arguments to the other canonical example of negative duties in the *Groundwork*: false-promising. Second, I must show how Hegel's emptiness *critique*, which includes his subjectivism objection, can be applied to the problem of natural actions. In doing so, I demonstrate that the Practical Contradiction Interpretation does not furnish a solution to the problem of natural actions, but to the contrary, reveals a fundamental subjectivity inextricable from the Moral standpoint. I turn first to the latter objective.

### **A. The Problem of Natural Actions**

While the Practical Contradiction Interpretation is of interest in and of itself, it is also of particular interest as the foundation to Korsgaard's proposed solution to the problem of natural actions. Following John Rawls, Korsgaard categorizes actions (for the purpose of



evaluating maxims) as either *conventional* or *natural*. Conventional actions are those actions that depend on a convention, or social practice, in order to be performed. For example, there must be the practice of promising before anyone can make a promise. The Practical Contradiction Interpretation is clearly applicable to conventional actions, or at least the standard examples of conventional actions, as we have already seen with the case of keeping a deposit.

In contrast, it is not immediately clear how the Practical Contradiction Interpretation can be extended to impermissible natural actions. Performing a natural action does not require a convention: one can run or swim, and importantly, assault or kill without the existence of a social convention. As such, the universalized maxim of an impermissible natural action does not lead to the annihilation of any convention and, correlatively, the means on which that action depends. As Korsgaard succinctly explains: “[w]hen an action's possibility depends only on the laws of nature it cannot become inconceivable through universal practice” (Korsgaard, 1996, 85). Impermissible natural actions, then, pose a peculiar problem for the Formula of Universal Law in that, no matter how wrong a natural action may be intuitively, it may nevertheless appear to pass the universalization test and consequently be judged permissible. The severity of this problem is represented by Dietrichson (1969, 188), whom Korsgaard takes as her interlocutor, and who provides the following example of a heinous yet ostensibly universalizable maxim: “If I give birth to a baby weighing less than six pounds, I shall do everything in my power to kill it.” As we shall see, Korsgaard does not think this maxim is properly formulated as it does not contain a specific intention or reason for killing the baby. We shall have to see, however, how successful Korsgaard’s proposed solution is.

The seriousness of this problem should not be underestimated as a mere technical failure. What the problem of natural actions means, in practical terms, is that the Formula of Universal Law fails to prohibit brute murder, rape or assault. This failure calls into question the entire universalization procedure as a source for determining particular duties. Solving the apparent problem of natural actions, then, is of utmost significance to the Kantian project.

Despite the independence of natural actions from conventions, Korsgaard argues that a solution is at hand and that a contradiction can still be derived from the universalization test if the relevant maxims of natural actions are formulated properly. Specifically, Korsgaard contends that we must include the specific intentions of the agent in the formulation of the agent's maxim.<sup>4</sup> To illustrate this solution Korsgaard asks us to consider two maxims: that of a thief, which properly formulated will include the intentions of “guaranteeing” his possession of property; and that of a murderer, which Korsgaard formulates so as to include the specific intention of advancing his career and “securing” a career position through killing his professional rival (Korsgaard, 1996, 98-99).

By including such intentions, the specific ends and goals of the agents are made explicit; and, Korsgaard argues, these ends and goals are either compatible or

---

<sup>4</sup> While Korsgaard is not this explicit regarding the specificity of intentions which must be incorporated into maxims, through a consideration of a selection of text I believe this attribution to be correct. First, Korsgaard writes that “If this interpretation is correct, then it is essential that in testing maxims of actions the purpose always be included in the formulation of the maxim.” (Korsgaard 1996, 92) Also, in her discussion of the stealing and murdering maxims Korsgaard refers to specific intentions. For example: “The idea here is that what the thief really wants is to make something his property, to have some *guarantee* that he will have it when he wants it” (98). I interpret “something” in this context to be some particular item and thus the purpose of stealing to be the acquisition of a specific and desired item when one wants it. Similarly, the purpose of the murderous colleague is to acquire a specific position, not to acquire any position, even a better position, in general. In general, then, the purpose is understood in terms of the specific end of the specific action.

incompatible with their own universalization:

So, if we include as part of the purpose that the agent wants to be secure in the possession of the end, we can get a practical contradiction in the universalization of violent methods....The method of dealing with natural acts which I have just suggested focuses on the question whether you could really achieve your purpose – with everything that purpose involves (i.e., security in its possession) in a world where your action was the universal method of achieving that purpose.  
(Korsgaard, 1996, 99)

With the inclusion of specific, explicit intentions, the agent's will becomes internally frustrated insofar as the end of *securing, guaranteeing or maintaining continued possession* of anything stolen or usurped through killing or thieving cannot be realized so long as killing and thieving *for these same ends* are universally practiced. The simple insight Korsgaard proposes is that if one's murderous maxim is universalized, one would likewise be murdered and hence the agent would not guarantee his position; likewise, if one's thieving maxim which includes securing possession of the stolen item were universalized, one could not be secure in his possession since everyone would use similar means to gain possession of what is now yours. In both cases a practical contradiction results only if these specific intentions are included in the maxim.<sup>5</sup>

Incorporating specific intentions, then, provides a solution to the problem of natural actions. However, I argue that Korsgaard's application of the Practical Contradiction Interpretation to the problem of natural actions trades one problem for

---

<sup>5</sup> It is worth noting that this response to the problem of natural actions is only necessary given the presupposition that there is no prior proscription concerning killing or similar natural actions. In other words, for the sake of this discussion it is assumed that the Formula of Universal Law must, independent of any other formulations of the Categorical Imperative, be capable of judging such natural actions as those described (self-interested killing and thievery) impermissible. This is a unique position, as other Kantians (see Wood 1999) argue that the Formula of Universal Law cannot and was not designed to function independent of the additional formulations. In short then, the need for this or any solution to the problem of natural actions may be peculiar to Korsgaard and others who share the view concerning the independence of the Formula of Universal Law.

another more significant problem. This problem may be understood as a variation of the Act-Description Problem. The Act-Description Problem concerns the ability of an agent to describe his actions in a variety of ways and thereby avoid the generation of a contradiction by the universalization procedure. As we shall see below, this problem is often concerned with what I term “gradational indeterminacy”; or the ability to add or remove details (e.g., day, time, place) from the description of the action. However, Hegel’s contention varies this objection, and contends that same action may be given different intentional descriptions concordant with the good the agent sees in his action. The application of Hegel’s objection to the problem of natural actions exemplifies his contention that the Formula of Universal Law results in moral subjectivity insofar as content, in the form of intentions, must be incorporated into the agent’s maxim in order to generate a contradiction. And yet, this necessary intentional content included in the formulation of maxims is the product of the agent’s heteronomous will. Thus, the intentional content constitutes imported heteronomous content that renders the moral permissibility or impermissibility of any given action unacceptably subjective.

## **B. Natural Actions and Moral Subjectivity**

How successful is this solution? Korsgaard herself seems to recognize that the solution has significant limitations along the lines of the Act-Description Problem. “Natural actions are not going to cease to exist if used wrongly, but their efficacy for some purposes may depend on their exceptional use. A great deal depends here on what the purpose is taken to be and how it is described” (Korsgaard, 1996, 97; emphasis added).

Korsgaard also concedes that violent acts where the end consists in the death or abuse of a person will not be solved by this approach: if one's end, for example, is simple revenge there is no form of security or similar end one is trying to achieve; all one wants is the death or harm of another.

Nonetheless, Korsgaard thinks that this solution is broadly successful. This success is due to the fact that security is a "special object of human interest." As we have seen, Korsgaard's solution depends on the incorporation of specific ends or intentions, and these intentions include the security or guarantee of the desired end. The murderous coworker desires not only to be promoted to the position, but to have security in this position; and the thief does not merely desire to steal property, but to guarantee the continued possession of the property once stolen. It is in this sense that Korsgaard, following Mill, contends that security is a "condition of the goodness of anything else" (Korsgaard, 1996, 99). Thus, Korsgaard concludes, "if we include as part of the purpose that the agent wants to be secure in the possession of the end, we can get a practical contradiction in the universalization of violent methods" (99). The incorporation of specific intentions that include the special interest of security is proposed as a sufficient condition for generating the requisite practical contradiction.

There is a three-part process in applying Hegel's moral subjectivity objection to Korsgaard's proposed solution. The first part consists in recognizing the simple idiosyncrasy of the ends adopted by agents, and the consequences this has for the broad success of Korsgaard's solution. The second part follows closely on the first by asking after the determination of ends. In other words, Hegel argues that whether or not the security is incorporated into the agent's maxim, and thereby the sufficient condition for

generating a practical contradiction, depends entirely on the convictions held by the subject. On Hegel's view, this determination of ends reveals an unavoidable and unacceptable moral subjectivity inherent to the universalization procedure. Lastly, an analysis of Hegel's objection will show that incorporating a desire for security is not itself a sufficient condition for generating the requisite contradiction; rather, the determination of what is essential to the action also determines whether a practical contradiction can be generated.

In considering the idiosyncrasy of ends, in light of Hegel's objection, it is helpful to start by considering Korsgaard's response to a similar concern which he thinks a Kantian needs to avoid:

Now, it may seem [when considering the murderer's maxim] that the purpose that is thwarted by universalization – that of staying alive – is not the same as the purpose in your maxim – that of securing the job. This would be bad. It is the fact that it is the purpose in the maxim that gets thwarted in the world of the universalized maxim that enables us to carry out the test without any extraneous information about the agent's desires and purposes. If it is some other, contingent, purpose that gets thwarted, then it looks as if the test (i) requires empirical information about what other purposes people have and (ii) functions idiosyncratically, giving different results to people with different desires. These are both conclusions the Kantian wants to avoid. We shall avoid them here by pointing out that this is not a case of an extraneous end being thwarted. Staying alive matters in this example because it is a necessary condition of having the job. (Korsgaard, 1996, 98)

The Kantian worry in this passage is remarkably close to Hegel's own critique. The crux of the concern is whether the purpose that gets thwarted has a contingent or essential relation to the explicit ends of the agent. In the examples given the purpose that gets thwarted is not identical to the explicit end of the agent, and thus its function within the universalization procedure is suspect. If the thwarted purpose were contingent (i.e., if the thwarted purpose were attached to the explicit end due to the desires and interests of the

agent, but was not a necessary condition of the explicit end) then the various possible desires agents may have for any given action must be investigated empirically. This empirical and heteronomous contingency would in turn result in the idiosyncratic function of the universalization test such that the impermissible maxim in question would only be impermissible for some agents with certain desires. Korsgaard thinks he can avoid this devastating result by showing that the desire to live is not “extraneous” to the agent’s explicit intention to secure a job, but is instead essential to the agent’s own internal goals insofar as one cannot secure a job if one does not remain alive.

We can characterize Hegel's charge of moral subjectivity as asking why the concern over idiosyncrasy should end with essentiality or contingency of the *thwarted* ends and not be extended to the explicit end; that is the intention of the agent ascribes to his action. In other words, Korsgaard's analysis is right: if the thwarted end is not essential to the explicit end of the maxim, then the universalization test will be idiosyncratic. In order to be essential *to* the explicit end, the thwarted end must be a necessary condition *of* the explicit end. Hegel's contribution to this analysis is to point out that the same essential relation between actions and explicit ends must obtain in order to avoid an idiosyncratic judgment of *actions*:

An injury done by a well-intentioned action whose determination is in other respects criminal and evil is, of course, also good, and all would seem to depend on which aspect of the action is the *most essential*. But this objective question is not applicable here – or rather, it is the subjectivity of consciousness itself whose decision alone constitutes the objective element. (*PR*, §140d)

In contending that there is no objective ground (at least available to Kant) for determining which aspect of an action is 'most essential', Hegel is first arguing that the determinate end of an action is not a necessary condition of the surd action – that is, the intention one

adopts in an action considered concretely is contingent. However, Hegel argues further that, within the context of *Moralität* the absence of an objective ground implies that what is essential to an action must then be determined by “the subjectivity of consciousness itself.” This constitutes an idiosyncrasy of explicit ends. To adopt the language of Korsgaard's analysis: Because the intention incorporated into a maxim is extraneous to the action itself, the permissibility or impermissibility of the action supervenes on this extraneous content. But, within the context of *Moralität*, this extraneous content is and can only be determined by the subject. And if one's duty is determined in terms of this extraneous content, then this idiosyncrasy implies that one's particular duty is inextricably subjective.

To reiterate some supplementary points from chapter 1, Hegel's contention is not that the universalization test does not function, but rather that the test functions idiosyncratically, giving different results to people with different subjectively determined intentions. Additionally, Hegel is not making the trivial point that, for any surd action there may be legitimate alternative maxims which will have to be tested by the universalization test. Two agents may perform the same surd action under different, even contrary, intentional descriptions; and the impermissibility or permissibility of this action depends on the description adopted. Hegel's objection does not concern the mere contingency of the intentional description, but rather the subjective origin of the description. Crucially, Hegel's objection applies to both maxims which pass the universalization test and those which do not. Thus, the significance of Hegel's argument may be put in the following way: First, for any surd action there may be illegitimate ends – ends an agent ought not to adopt – which nevertheless can be heteronomously adopted



and pass the universalization test; and second, the legitimate ends which are adopted are equally a product of the agent's own arbitrary will. In either case, intentional idiosyncrasy results in moral subjectivity.

To make Hegel's point more concretely, we may consider the following modifications to Korsgaard's examples.

*Well-intentioned Murderer.* Two individuals are competing for a promotion. The first candidate is more qualified than the second; but the second candidate is more charming. The first candidate assesses the situation and realizes that he, though the more qualified candidate, will likely lose to the less qualified but charming rival. With a sincere interest in protecting and improving the company, the first and more qualified agent decides to kill the less qualified candidate and to secure the position for himself.<sup>6</sup>

*Hobbyist Burglar.* According to Korsgaard, in order to properly formulate the thief's maxim, we must include the intention to guarantee continued possession of what is stolen. This ascribes a subjectively idiosyncratic intention to the agent such that the good of thieving consists in the acquisition of the item desired. The Hobbyist Burglar does not share this end; for him the good end is intrinsic to burglary itself. His intention is to pull off the grandest and most elaborate burglaries for the sheer thrill, enjoyment and appreciation of the skill and technicality involved – just as many other hobbyists do. Guaranteeing property is of secondary or even no consequence.

These scenarios represent two different approaches of applying Hegel's objection and showing the idiosyncrasy of intentions. In the case of the *Well-intentioned Murderer* security is an integral part of the agent's end in that securing the position for himself is a necessary condition for achieving his end of protecting and improving the company. However, though security is integral to the end, and even a necessary condition of achieving his end, this integration is not a sufficient condition for generating the requisite contradiction. Universalization fails to generate a contradiction because securing the

---

<sup>6</sup> Securing the position is part of this agent's end, but it is not the governing end. That is, the agent does not kill in order to secure; he kills in order to secure in order to protect and benefit the company. It is due to this adopted end that his maxim can be universalized, because it is only in cases where the company is at risk that murder takes place.

position for himself is not (what is perceived as) essential to the action. What is essential, from the standpoint of the agent, is the success of the company, and this determination of what is essential introduces a limit on universalization.<sup>7</sup> In contrast, the *Hobbyist Burglar* does not integrate security into his maxim because securing the property is of no interest to the burglar. Of course, security still remains a background necessary condition on the enjoyment of his hobby; but this condition obtains for any hobby or activity (one needs a general sense of security to enjoy playing racquetball at the gym). What matters in the context of universalization is whether security is integral to the end itself (e.g., securing property), and not merely that security is in general and indirectly “a special object of human interest.” Security, then, is neither a necessary component of an agent’s end, nor a sufficient condition when incorporated into the agent’s end.<sup>8</sup>

It would be mistaken to analyze these examples truistically in terms of possible different intentions. To the contrary, the shared ethical significance of the *Well-*

---

<sup>7</sup> This example is of further interest because the agent, we may stipulate, knows what is morally relevant, knows that killing violates a generally valid moral precept, and applies the universalization test with a genuine interest in determining whether or not his case is justifiably exceptional (see Herman, 1993e, 77-78). Even in this morally rich context, the universalization test is insufficient for proscribing the impermissible.

<sup>8</sup> One might object by questioning whether the illegitimate maxims of the *Well-intentioned Murderer* and the *Hobbyist Burglar* are universalizable. I believe they are. In the case of *Well-intentioned Murderer*, the agent’s end of protecting and improving the company from bad consequences of promoting unqualified candidates is not thwarted by the universal practice of killing unqualified candidates likely to win promotions for the wrong reasons. To the contrary, if anything the agent’s end is better realized by the universalization of his maxim. If incompetent but charming candidates were always killed, then the company would be protected *either* by their death *or* by the unwillingness of similar candidates to pursue promotion (cf. Korsgaard’s discussion of universalizing the maxim “Help the poor” [1996, 86-87]). Indeed, this analysis gives a different perspective to Korsgaard’s own contention that “whatever it is about the old selection process that makes [murder] possible would be changed” (100). Korsgaard assumes here that the convention would be changed to prevent murder by, for example, not making the candidates public. However, it is just as rational to institute the killing or threat of killing poseur candidates to ensure the success of the company. Moreover, either modification to the convention of selecting candidates can be understood as concordant with and conducive to security. Likewise, while it may become more difficult to perform burglaries if the hobby were universalized, there is no contradiction in universalizing this hobby, no more so than universalizing any number of other hobbies (e.g., stamp collecting or antiques).

*intentioned Murderer* and the *Hobbyist Burglar* is that they both perceive an illegitimate end as *essential to* and thus as *the good* in the situation. The *Well-intentioned Murderer* perceives the success of the company as superseding the life of an incompetent but charming competitor. Likewise, the *Hobbyist Burglar* perceives the thrill and requisite expertise of burglary as superior to the property rights of his victims. According to Hegel, the subjective prioritization of such possible illegitimate intentions is a general property of formulating all maxims: “But that subjective determination which I give to the good is the good which I *know* in the situation, i.e., my *good intention*” (PR, §140d). Moreover, the ends we know are the ends adopted, learned, habituated or otherwise acquired *heteronomously*. In this way, the heteronomous origin of peculiar ends undermines the rational foundation to the universalization procedure: if the end we perceive as good is the end we know in the situation (and thus an end adopted heteronomously), then the impermissibility of the murderer's or thief's maxim – even of those agents who wish to secure or guarantee their ill got gains – is attributable to their own idiosyncrasies. This subjective idiosyncrasy is inextricable from the universalization procedure.

The severity of this conclusion is, perhaps, best brought out by the fact that Korsgaard has not provided a solution, consisting in a reformulation of the maxim, to Dietrichson's original challenge of a mother who adopts the maxim “If I give birth to a baby weighing less than six pounds, I shall do everything in my power to kill it.” What specific intention must we attribute to a mother who intends to kill a baby over a certain weight? It would not seem that a particular and relevant intention of security applies to an apparently arbitrary maxim. The intention of the mother no doubt seems alien (or atrocious) to us; however, this only further illustrates the possibility of deep idiosyncrasy

and subjectivism which the universalization of maxims of natural actions cannot avoid.

One might object in the following way: Hegel has (again) misunderstood the Kantian objective. The aim of the Formula of Universal Law is to evaluate an agent's proposed maxim, and thereby the *moral worth of the agent*; yet Hegel's objection concerns the possible ascription of impermissible maxims to a given concrete action. In this way, Hegel's objection is misplaced.

In one sense this objection is right: Hegel has changed the focus of the dialectic. But this is not due to a misunderstanding of Kant's objective. Rather, Hegel is posing an external objection to *Moralität*. Hegel's point is that the Kantian, by means of the Formula of Universal Law, cannot rule out impermissible actions – like murdering one's boss or killing one's child – unless the right intention is ascribed to this act. And yet, which intention gets ascribed to a surd action (within the context of Kantian *Moralität*) is and must be determined by each individual subject. This is unacceptable to Hegel. Does it constitute an internal objection to Kant? No, insofar as the Formula of Universal Law does function given the requisite intentional content. And yet it is a powerful *external* objection in that despite the successful functioning of the Formula of Universal Law given the right intentions, it still leaves room for permitting killing, rape, assault and abuse of myriad forms. And further, even when a natural act of abuse or violence is ruled impermissible, this success relies on the subjective determination of what is good in the action.

However, in another sense Hegel has not changed the focus of the dialectic because his analysis concerning the idiosyncratic ascription of maxims is relevant to the *value* of actions, and thus to the moral worth of agents:

But if a good heart, good intentions, and subjective conviction are said to be the factors which give actions their value, there is no longer any hypocrisy or evil at all; for a person is able to transform whatever he does into something good by the reflection of good intentions and motives, and the element of his *conviction* renders it good.... My good intention in my action and my conviction of its goodness *make it good*. (PR, §140e)

On the Kantian view, the moral worth of an agent is determined by whether the agent acts from duty. An agent acts from duty if he abides by the Categorical Imperative; that is, if he does what is permissible and refrains from doing what is impermissible *because* such acts are accordant with or contrary to duty, respectively (see Kant's *Groundwork* [4:403]). However, Hegel has shown that the permissibility or impermissibility of a maxim is also determined by what the agent perceives as essential and good in the action and the incorporation of this essential element into the maxim. Thus, an agent may perform any action and judge it permissible in terms of the intention he ascribes to the action. Furthermore, he may perform any action from duty, including those actions widely judged as immoral. Which is to say that an agent may steal, abuse and kill and judge all these actions as permissible in terms of a subjectively imputed intention. If this is correct, then Kantians would be committed to attributing moral worth to an agent, no matter what concrete action he has performed, so long as he perceives a good in his action, adopts this good as his intention, and acts from (a sincere sense of) duty.<sup>9</sup>

---

<sup>9</sup> The alternative examples considered above may help illustrate the subjectivity of moral worth. Let us further stipulate that the *Well-intentioned Murderer* sincerely considers whether he may act on his proposed maxim, and submits his maxim to the universalization test. It is only after and because the maxim passes the universalization test that he adopts it as his actual maxim. Which is to say, the agent acts for a good intention *and* because the action conforms to his duty. Furthermore, the agent sees this killing as a selfless act of dutiful gratitude and loyalty to his company. After all, though he is securing a position, he is also taking on additional responsibilities and is assuming the risks that accompany murder (e.g., prosecution, imprisonment, ostracism) to ensure the prosperity of his company. In every respect, then, the *Well-intentioned Murderer* is not only in accordance with duty, but also of moral worth. The sufficient condition for this attribution of moral worth is the agent acting from duty and the agent's *subjective conviction* that he is essentially acting for a good reason.

To be clear, the determination of moral worth is not formally subjective, insofar as the criterion for moral worth remains acting from duty. What is subjective is the ground for determining whether a particular agent is morally worthy. The determination of a particular agent with a particular intention as morally worthy depends on the prior determination by the agent himself of what is essential and good in a given action. This is to say that while moral worth in the abstract remains objective, the moral worth of any particular agent always and necessarily has a heteronomous foundation. This follows the general form of Hegel's critique: *Moralität* can only provide us with the form of moral worth; actual moral worth necessarily involves the heteronomous will of the agent, and is thus inextricably and – for Hegel – unacceptably subjective.

### **C. Conventional Actions: False-Promising**

As we have seen, the foregoing critique in terms of moral subjectivity is external in that it does not target the function of the universalization test itself, but rather the tenability of the test as source for determining our universal and objective particular duty. In contrast, Hegel's emptiness objection targets the functioning of the universalization procedure itself by demonstrating that the universalization does not, and cannot, produce substantive results. This internal objection also contends that heteronomous content incorporated into the agent's will plays a necessary role in generating a contradiction. Unlike natural actions, this content consists in a normative specification which is implicitly invoked and attributed conceptual necessity in the universalization process. Hegel contends that the incorporation of this content enables the universalization test to function, but at the cost of substantive results. More specifically, he contends that the

conclusion of the test, when considered within the context of the normative specification assumed, consists in a tautology, and this tautological vacuity renders the universalization test itself superfluous. This objection has been developed in chapter 1 in terms of the deposit scenario; the current objective is to show how this objection generalizes by applying Hegel's arguments to the false-promising scenario.

Most readers will be familiar with the scenario: An agent in financial need considers whether he can make a false promise to an acquaintance who may lend him money. His maxim as formulated by Kant is "when I believe myself to be in need of money I shall borrow money and promise to repay it, even though I know that this will never happen" (*G*, 4:422). We ask if this maxim can be universalized and, according to the Practical Contradiction Interpretation, consider whether the means to the agent's end specified in his maxim would be undermined in the world where that maxim is universally followed. The maxim is rejected as impermissible, according to the Practical Contradiction Interpretation, because the universalization of the false-promising maxim would annihilate the convention of promising (as we know it) and, thus, the means to the agent's end.

Hegel does not reject this conclusion, but contends, rather, that this conclusion relies on specific normative content incorporated into the agent's maxim and will. This specific normative content is contingent, in that there exist opposing possible specifications of norms which may be incorporated into the agent's will, but only one of these opposed specifications has been posited within the will of the agent. On Hegel's view, this constitutes an unwarranted and *heteronomous* privileging of one normative specification over another. Hegel argues further that if this content were not posited, a

contradiction could not be generated. And lastly, this determinate content is of such a thick normative specification that the resultant contradiction, which is supposed to inform us of our duty, is tautologous: if promising is to be then promising must be.

To clarify Hegel's objection with regards to false-promising, we may make explicit what is implicit within the false-promising scenario. When and under what conditions is a false-promising maxim impermissible? The answer appears obvious from the description of the case: when the agent intends to *exploit* the convention necessary for his own selfish benefit. However, this exploitation is not part of the description of the case. Rather, it is a desideratum of the universalization test to *show* that the maxim is exploitative or exceptional in character. Descriptively, the agent considers only whether or not he may "borrow money and promise to repay it even though I know this will never happen." What makes this description so evidently exploitative?

An answer requires identifying what else must be assumed when considering the description. The first assumption is how to understand "promising." For example, the performance cannot be understood as a mere linguistic act – an utterance that money will be paid back – as there are any number of ways in which such an utterance may be made without deception or exploitation. Rather, from his own explication and the conclusion reached, it is clear that Kant means for the reader to understand the act of promising as an act of establishing trust. It is only by establishing trust that the agent can deceive his associate, and thereby exploit the convention. Yet establishing trust relies on expectation on the part of the lender that the money will be paid back; and deception, in turn, requires knowledge on the part of the agent that this is what is expected. In other words, both the agent and his associate must have a practical facility with the normative context. Each



must understand that the utterance, in this context, is to be taken literally and that by this utterance the agent is morally bound to repayment. Said differently, a necessary condition on the success of the example is that both agents understand that deliberate failure to repay the loan, without good reason, would be *wrong*.

Such explicating of what is implicit is in one sense trivial. The analysis simply articulates that the agent only makes a promise if there is a norm of promising as that institution is generally and normally understood. But what it means for there to be a norm of promising as that institution is generally and normally understood is for there to be a *proscription against falsely giving a promise when in need*. In other words, the deception in *this* context and for *this* reason is only possible because it is *proscribed* in these particular conditions. The interesting point is that we knew this before the universalization procedure, and we *had to know this* before considering the maxim in order to properly and successfully evaluate the maxim. Hegel's contention is simply that this prerequisite condition for the possibility of exploitation renders the test trivial, insofar as false promising is impermissible only if it has been antecedently posited as impermissible within a normative specification assumed and incorporated within the agent's will. By making this explicit, the conclusion of the test is reduced to a superfluous tautology: One may not make false promises when one is in dire financial need *only if* false promising for this purpose and under these conditions is impermissible.

As in chapter 1 we may illustrate Hegel's position by developing his claim that *if* this particular normative specification of promising were not posited, then the target contradiction could not be generated. To develop an opposing normative specification we may consider a hypothetical evolution of norms within the convention of promising.

Consider first a familiar normative deviation within our own convention of promising: a promise is made to pay back a loan, but following this promise there is expressed the absence of any expectation that the loan will be paid back. An agent facing a financial crisis asks his friend for a loan; the friend loans the money and the agent promises to repay him when he can. In a situation like this it would not be uncommon or unexpected for a friend to respond, "It's not a loan; it's a gift." For some, at least, this is not merely a matter of generosity: in this context of dire need and friendship, the right thing to do is to offer to pay the money back *and* for the friend loaning the money to refuse the promise. Some consider a loan, conventionally understood and under these conditions, to be inappropriate between friends.

Next, imagine that this normative specification evolves so that when a promise is given to repay a loan under the same circumstances of dire need and friendship, there is not any expectation that it be honored. The imagined norm is a codification of the normative understanding which is expressed between friends in the original scenario; the difference is now that this norm is understood and practiced without the explicit annulment of the repayment obligation. The norm, for example, might dictate that he who asks for money, out of respect and to avoid embarrassment, ought to promise to pay the money back; but the norm further dictates that the lender not expect any repayment – it would be unthinkable, morally, to expect repayment from a friend in such a dire financial situation and equally unthinkable to embarrass his friend by refusing the offer of repayment.

Could an agent in dire financial need universalize his maxim to make a promise to a friend with no intention of paying it back given this normative specification? Because

of the highly specific context of making a false-promise to a friend when in dire financial need, the maxim can be universalized. The specificity of the context produces what we – as outside observers – would ordinarily think of as a false-negative: the universalization procedure does not rule a maxim impermissible when it should. We would ordinarily consider this a false-negative because within our assumed normative specification of promising there does not obtain an understanding between agents that a promise made is not expected to be honored. However, from inside the normative specification – i.e., to those with an understanding of the norms – no false-negative has been produced: the universalization procedure has functioned just as it should even though a false-promising maxim between friends was conceived as a universal law. In other words, the functioning of the universalization procedure conformed to the normative expectation posited by the agent considering the proposed maxim.

To illustrate further this conformity of the universalization procedure to the posited normative specification, note that the agent *cannot* universalize a maxim where he pretends to be in dire financial need and to make a false-promise with no intention to pay it back when there is the expectation that the promise will not be honored. Universalizing this maxim would mean that a friend would no longer believe an agent when he makes a false-promise because he is in dire need, because everyone pretends to be in dire need, so the agent's expression that they are in actual financial need would be ridiculed.

Simply put, the agent cannot universalize a maxim which violates a proscription of the normative specification assumed. When a standard convention of promising is posited, the agent cannot will a false-promising maxim. With the modification to

promising assumed, the agent can universalize a false-promising maxim (i.e., a maxim to affirm his intention to pay back the money with no intention of doing so) because the normative specification permits false-promising of just this form. The agent cannot, however, will a maxim that violates the norms of this modified convention.

Now consider a final evolution of this opposing convention where this exception to the norm of promising is generalized. We might call this convention *Honor promising*. In this convention, false-promises are made as a form of avoiding embarrassment when one is in a dire financial position. The false-promise is made to someone – friend or otherwise – who can help the agent out of the financial crisis. As in the above scenario of promising between friends, the false-promise is accepted without any expectation that the money be repaid. To expect repayment from someone in such dire a financial condition would be shameful. Additionally, let us stipulate that it would shame the *lender* if the agent who borrows money would actually pay the lender back. That is, it would bring shame on the lender if the agent were to honor his promise. In this normative context, false-promising to repay a loan when faced with financial hardship is *prescribed*.

Could an agent universalize their maxim to give a false-promise? Since the maxim is a single instantiation of the norm, to contend that the agent could not universalize this maxim would require denying the possibility of the convention – that *Honor-promising* is somehow inconceivable. There is, however, nothing inconceivable about the convention itself, and so the maxim must be considered universalizable.

As in the above example of promising between friends, the functioning of the universalization procedure conforms to the normative expectations of *Honor-promising*. However, there is a further implication of the *Honor-promising* convention; namely, that

an agent who considers a maxim to honor a promise could not universalize his maxim. To clarify: Imagine an agent who, motivated by self-interest, considers a maxim where he promises to repay someone who lends him money *with* the intention of paying it back. The purpose of this honest promise is, perhaps, to embarrass the lender or self-aggrandizement. If the agent were to universalize his maxim, then everyone would honor their promises and, accordingly, the institution of *Honor-promising* and, thereby, the means to the agent's end would be annihilated. Thus, when the normative specification of *Honor-promising* is posited, honesty in the form of an authentic promise when one is in financial need is shown to be impermissible by means of the universalization procedure.

This progressive modification of the promising convention is a concrete example of Hegel's counterfactual contention that, while the universalization procedure functions, this functioning is constrained by the normative specification in which the test is embedded. What Hegel means to show is that a contradiction must be a contradiction with a presumed content, and precisely "because the criterion is a tautology, and indifferent to the content, one content is just as acceptable to it as its opposite" (*PH*, §430). This is to say that the universalization procedure functions, but only to show a formal relation between two contents: the content of the posited normative specification and the content of the agent's maxim. The universalization test, thus, merely reveals that there obtains a contradiction endemic to a particular ethical life (*Sittlichkeit*). However, that a contradiction obtains, and the universalization procedure can reveal this formal relation, is true of any number of *Sittlichkeiten*.

In sum, "[l]aws are no longer given, but *tested*; and for the consciousness which tests them they are *already* given" (*PH*, §429). The *Sittlichkeit* provides the laws, and

thereby the normative grounds in terms of which the universalization procedure will or will not function. And the agent, in testing his maxim, only discovers what is possible or impossible within a presumed normative convention.

In chapter 1 I considered an important objection to this analysis which is worth briefly revisiting. This objection rebuts that the universalization test is not designed to demonstrate that the particular deception of the false-promising scenario is impermissible, but rather, that this deception consists in exploitation for individual benefit, and thus the proposed maxim exemplifies a general *form* that is impermissible. In other words, the test is not meant to demonstrate *that* false promising is contrary to duty, but *how* false promising is contrary to duty. Korsgaard puts it in the following way, “Intuitively speaking, the test reveals unfairness, deception, and cheating” (1996, 92). More formally, the test identifies those maxims which are *exceptional in character*.

As we saw in chapter 1, this rebuttal merely pushes the challenge back to a new level of abstraction, and may be answered with the modified question “When and under what conditions is the agent's maxim exceptional in character?” Whether or not a given maxim is exceptional in character is conditional on the normative convention in which the maxim is considered. In other words, Hegel makes no objection to the proscription against exceptional maxims *per se*; rather, Hegel's objection targets the conditionality of this abstract proscription in relation to the determination of our actual duties. Said differently, Hegel's contention is that those maxims which are exceptional in character relative to one ethical practice may be incorporated into an opposing normative specification as unexceptional principles of action; that is, permissible exceptions within the ethical practice. The implication of the relativity of exceptional maxims is that any

action whatsoever may be permitted or even required so long as an exception for the principle of the action is incorporated into the normative specification of the relevant ethical life. Indeed, as we saw in the case of *Honor-promising*, the proscribed maxim in the false-promising scenario may be permissible and even prescribed given the right normative specification. Thus, identifying the form of duty is not sufficient for ensuring the proscription of any evil, since this very evil may be codified as good.

If this analysis is right, it is devastating. Essential to Kant's project is establishing universal duty which transcends what has been or can be established conventionally. While it is true that the form of impermissible maxims transcends any particular normative convention, particular duty does not. Our particular duty is necessarily indexed to a particular normative specification. Thus, the impermissibility of false-promising is indexed to a convention of promising where false-promising of just this kind is impermissible; and conversely, telling a false-hood when in need may be permissible when and if the normative specification makes allowances for false-hoods under the relevant particular condition of "need." And this indexing generalizes: any conventional action whatsoever may be permissible or impermissible according to normative specification of one's convention. In this way, conventional norms are essential to the determination of particular duty in such a way that particular duty is relative to one's convention. It is in this context that we should understand Hegel when he writes that "the sole thing at issue is what lies outside the capacity of this practical legislation of pure reason, namely to decide which of the opposed specific things must be lawful" (*NL*, 78). Insofar as our particular duties are conditional on the assumed normative specification, and the positing of the relevant normative specification is done independently of pure

practical reason, the positing of our normative conventions independently of pure practical reason must be central to our ethical theorizing. We must decide independently of the universalization procedure which normative specification will be incorporated into the universalization procedure.

## II. MAXIMS AS GOVERNING PRINCIPLES

As stated in the Introduction, there is no general consensus regarding how maxims should be understood. In the preceding section, maxims were formulated to include specific intentions, such as “securing property” or “guaranteeing a position,” and constituted principles for particular proposed actions – for example, killing a rival. I refer to this interpretation as a “thick” interpretation of maxims. Onora O’Neill, in her collection of essays *Constructions of Reason*, rejects the thick interpretation of maxims and argues that we must construe maxims as governing or fundamental principles of action *sans* specific intentions.<sup>10</sup> I will refer to this view of maxims as a “thin” interpretation of maxims. This interpretation is of particular importance, because O’Neill’s express desideratum in

---

<sup>10</sup> I focus exclusively on O’Neill’s position as presented in *Constructions of Reason*. Her earlier work, *Acting on Principle*, addresses similar issues (i.e., the description and formulation of maxims). In this earlier work, O’Neill (Neill) argues that “Kant’s solution to the problem of relevant descriptions is to claim that any voluntary act has a maxim to which the Categorical Imperative should be applied” (1975, 41). This inclusivity, likely because of the problem of false-positives and false-negatives, was revised in *Constructions of Reason*.

As with interpreting Korsgaard, I use both O’Neill’s explicit characterization of maxims as well as her examples in attributing this view of thin maxims to her. O’Neill clearly states that “An agent’s maxim in a given act cannot, then, be equated simply with intentions” and continues, characterizing maxims in terms of “underlying principles or intentions” (1989, 84). O’Neill’s examples, such as the maxims of slave-holding or abrogating one’s duty, further clarify what is and what is not an underlying principle. At times her construal of maxims may appear specific, such as her example of a maxim to “make a visitor feel welcome.” However, even in the case of hospitality, I understand the maxim as incorporating an intention to make *any* visitor feel welcome, and not *this particular* visitor – otherwise the maxim would encounter the same difficulty of excessive specificity Wood has observed (see 79f, below).



developing this position is to demonstrate how universalization tests generally, and specifically the Formula of Universal Law, can guide action without incorporating illicit heteronomous content.<sup>11</sup> The thin construal of maxims thus serves as a solution to the Act-Description Problem: by establishing a criterion for construing maxims that eliminates extraneous material, a maxim at the appropriate level of determinacy may be imputed to the agent (and by the agent) to be universalized.

It must be noted that, while there is disagreement regarding precisely how maxim ought to be construed, there are some necessary conditions. Freyenhagen (2012, 62) identifies four such conditions: maxims appropriate for testing must be (1) from the first-personal perspective; (2) “minimally rational, where this means that they must conform to the hypothetical imperatives in question”; and (3) the maxim must be the actual, sincerely held maxims of the agent. In the examples of maxims given below these conditions are maintained. Freyenhagen also contends that, in formulating maxims, we are concerned with “action types” and not specific acts. Unfortunately, his contrast examples fail to identify the different possible grades of description; he writes, “we are interested in, for example, deception for personal gain, not the specific time and day this takes place” (*ibid.*). While I agree that the latter example is universally considered too specific, the former may, as we have seen above, be considered too general. Nonetheless, this condition accords well with O’Neill’s thin construal of maxims and will thus be

---

<sup>11</sup> O’Neill is concerned with rebutting what she terms “heteronomous readings” of Kant. “Nevertheless, Kant interpretation, particularly in English, is rich in heteronomous readings of the Formula of Universal Law and in allegations that (despite claims to the contrary) it is impossible to derive nontrivial, action-guiding applications of the Categorical Imperative without introducing heteronomous considerations” (O’Neill, 1989, 82). Interestingly, in a note to this passage O’Neill references several authors, including Schopenhauer and J. S. Mill, but makes no mention of Hegel. However, in the Introduction to the second edition of *Acting on Principle* (2014) O’Neill does recognize “German Idealism” as presenting a challenge for how formal principles could be action guiding.

maintained for the examples in this section.

With these considerations in mind, the application of Hegel's objections must be adapted to O'Neill's interpretation. This is specifically true for Hegel's emptiness objection. Hegel's emptiness objection, as I have argued, contends that the successful function of the universalization test relies on the heteronomous positing of a particular and contingent normative specification, in terms of which the normative concepts in the will and maxim of the agent are understood. As I explicate below, on O'Neill's interpretation the consideration of maxims is removed from any specific normative specification. Concretely, one does not consider whether an agent can make a false-promise; but rather, considers whether, in general, one can adopt a fundamental principle of deceit to govern their actions. As such, Hegel's emptiness critique cannot be applied directly to cases where maxims are construed thinly. However, this is not the case for Hegel's moral subjectivity objection, which is directly relevant to maxims *qua* governing principles.

With these objectives in mind, this section is organized as follows: First, I explicate O'Neill's position concerning maxims as governing principles, including the putative advantages and remaining challenges of this interpretation. Specifically, I argue that the objective of circumventing false-positives and false-negatives by formulating maxims thinly is not achieved if we are consistent in our formulation of maxims. Second, I argue that a thin interpretation of maxims does not solve the problem of moral subjectivity. I situate this now familiar objection in relation to the problem of multiple maxims – a problem O'Neill explicitly claims is circumvented by a thin interpretation of maxims. Lastly, I argue that a thin interpretation of maxims faces a unique problem

concerning the permissibility or impermissibility of ancillary principles – that is, those principles of high intentional specificity which are in service to and help realize the general end of the governing principle of action. The problem of ancillary principles reintroduces the relevance of Hegel's emptiness critique by presenting a dilemma: If ancillary principles must pass the universalization test, then the universalization test cannot avoid the problem of false-positives. However, if it is not the case that ancillary principles must pass the universalization test, then all kinds of obviously impermissible actions may be performed in service of worthy maxims.

### **A. Governing Principles and Gradational Indeterminacy**

The crucial attribute of a thin interpretation of maxims is that maxims are governing principles which guide and make intelligible the adoption of ancillary principles, and do not include specific intentions. “*Maxims are those underlying principles or intentions by which we guide and control our more specific intentions*” (O'Neill, 1989, 84). O'Neill elucidates this construal of maxims with an example of *hospitality*:

For example, in making a new visitor feel welcome I may offer and make him or her some coffee. As I do so, there will be innumerable aspects of my action that are intentional – the choice of mug, the addition of milk, the stirring – and there will also be numerous aspects of action that are “below the level of intention” – the gesture with which I hand the cup, the precise number of stirs and so on. But the various specific intentions with which I orchestrate the offer and preparation of coffee are all ancillary to an underlying principle. (84)

There is a sense in which this understanding of maxims is perfectly clear: the maxim is the principle of action which explains the more specific actions and intentions, and is thus the grounding reason for why we are performing more specific, particular actions. I would not be stirring milk into coffee, for example, if it were not for my intention to

‘make a new visitor feel welcome.’ However, there is another sense in which the idea of maxim *qua* governing principle is opaque: this is because we can formulate maxims at varying degrees of generality or specificity. Thus, the question remains, “At what level of generality or specificity must we formulate our maxims, and how do we determine the proper degree of generality of specificity?” Simple examples can be seen with O’Neill’s own case of hospitality. We may formulate the maxim of the hospitable agent as ‘making a new visitor feel welcome,’ though we also may take this principle of hospitality as ancillary to the even more general principle of ‘abiding by traditional standards of etiquette.’ Likewise, the agent may conceive his of action, and thus formulate his maxim, in terms of ‘making a *neighbor* feel welcome.’ Any of these formulations could serve as an explanation for ancillary intentions.

While this gradational indeterminacy may seem trivial, resolving it is absolutely critical to the success of the Formula of Universal Law. As Allen Wood (1999) has made clear, the possibility of different grades of maxim formulation poses the dual problem of false-positives and false-negatives.<sup>12</sup> On the one hand, if I formulate my false-promising maxim at a greater level of specificity – for example, willing to borrow money from my gullible friend Larry before the end of the year – then my maxim may be universalized without contradiction. This is clearly a false-negative: an impermissible intention of

---

<sup>12</sup> It should be noted that the use of “false-positive” and “false-negative” in the Kantian literature is peculiar, and departs from common usage within, for example, the scientific and medical fields. Standardly, a false-positive is when a test indicates that a particular condition is present when the condition is, in fact, not present. In the context of the universalization procedure, a false-positive does not falsely indicate the presence of a condition, but rather indicates the actual presence (or generation) of a contradiction, and thus a proscription, contrary to what is thought should be. Thus, for example, a false-positive occurs when the universalization procedure rules that one cannot act on an innocent maxim, such as “play tennis on Sunday.” Likewise, a false-negative occurs when no contradiction is generated for intuitively opprobrious maxims.

false-promising passes the universalization test. Likewise, I can generate false-positives by incorporating specific aspects of the situation into my maxim. Wood's provides a helpful example of playing tennis: If I consider my maxim to play tennis at the local court on Sunday, when everyone else is at Church, I clearly cannot will this to be universal law; doing so would undermine the means (an available tennis court) to my apparently permissible end. Hence, innocuous maxims formulated at greater levels of specificity may be falsely ruled impermissible by the universalization test.<sup>13</sup>

Wood suggests that the problem of gradational indeterminacy may be the reason why O'Neill developed her position of maxims as governing principles. However, he further contends that this proposal does not solve the problem insofar as the position "still doesn't tell us what is 'underlying', that is, how deeply we should probe into an agent's intentions in order to find what is morally relevant" (Wood, 1999, 104).

O'Neill does provide a criterion for determining what is "underlying" worth exploring – if only to understand precisely why her interpretation is ultimately unsatisfying. O'Neill thinks the grade of generality or specificity should be determined by the dependency that obtains between the action and different elements of the maxim:

For principles of action need only incorporate *some* true description of an agent and *some* true description of the act and the situation.... But an agent's maxim in a given act must incorporate just those descriptions of the agent, the act and the situation upon which the doing of the act depends. (O'Neill, 1989, 84)

---

<sup>13</sup> It is worth noting that the very idea of a false-positive or false-negative in the context of the universalization procedure highlights Hegel's general point regarding the necessity of external and illicit content. For, we may ask, how is one to determine if the universalization procedure has generated a false-positive or false-negative without the aid of some further moral standpoint? Furthermore, fine-tuning the formulation of maxims so as to avoid false-positives and false-negatives requires access to a moral standpoint or ethical practice that is normatively prior to the Formula of Universal Law itself. As such, if we were to solve the problem of false-positives and false-negatives such a solution would either be in terms of a normative specification we antecedently accept or would be entirely independent of successfully avoiding false results.

This criterion for formulating maxims has some intuitive appeal. Consider again O'Neill's hospitality maxim: 'making a new visitor feel welcome.' This maxim contains all and only those elements needed to explain and govern the adoption of specific intentions and actions that actualize hospitality. Though one does many specific things in being hospitable, none of these particulars are necessary for explaining *why* one is performing those particular actions, and are thereby superfluous to the formulation of the governing principle. Hence, the kind of dependency O'Neill seems to be after is a reasons-giving dependency: a maxim must incorporate those elements which make intelligible the general act (being hospitable) as well as the complementary actions and ancillary principles.

In order to evaluate the success of this dependency criterion we must see if it generalizes to other cases. It does not, and indeed O'Neill herself formulates exemplifying maxims with diverse elements and at varying degrees of specification. In this way, the apparent success of the universalization test depends on an inconsistent formulation of maxims. Moreover, a consistent application of the dependency criterion for formulating maxims encounters the same problem of false-positives and false-negatives Wood enumerates.

Consider, for example, O'Neill's analysis of maxims of slavery and slave-holding. O'Neill contends that the maxims to 'become a slave' or, alternatively, to 'become a slaveholder' both generate a contradiction of conception: "But if everybody becomes a slave, there would be nobody with property rights, hence no slaveholders, and hence nobody could become a slave." Likewise, the maxim to become a slaveholder, when universalized, leads to a contradiction of conception in that "everybody would have some

property rights, hence nobody could be a slave; hence there could be no slaveholders” (O’Neill, 1989, 96). The problem with these examples is that the reason one has in choosing to become a slave or to be a slave-holder does not (or need not) end with slavery itself, which means that there is no maxim of slavery or slave-holding to be tested. In other words, O’Neill has not followed her own dependency criterion: the reason one has to adopt a slavery or slave-holding maxim depends on reasons distinct from ‘becoming a slave’ or ‘owning a slave’ per se. Consequently, if we identify the underlying reason on which the slavery or slave-holding maxims depend, we see that they are innocuous. Thus, rather than exemplifying the effectiveness of maxims *qua* governing principles, these pseudo-maxims exemplify the danger and challenge of determining the right level of specificity or generality when formulating maxims.

To see this further, consider that the gratuitous specificity of the slavery maxim is evident from the fact that the action and intention of voluntarily becoming a slave can easily be made intelligible in terms of an alternative, evidently universalizable maxim; namely, “effectively labor or work for my living with whatever means or opportunities are available.” This formulation follows O’Neill’s own dependency criterion which dictates that we omit any superfluous specificity.<sup>14</sup> Consequently, on O’Neill’s own view, we *must* formulate the slavery and slave-holder maxims such that only the elements on which the action depends are included. But this implies formulating the slavery and slave-holding maxims (as well as other possibly impermissible maxims, such as a maxim of prostitution) in terms of the governing intention to effectively earn a living with the means available to me. Thus, the dependency criterion actually allows for maxims which

---

<sup>14</sup> The same may be said for Freyenhagen’s condition that maxims be formed for “action types” such as “deception for personal gain” (2012, 62).

pass the universalization test and permit slavery and slave-holding – that is, maxims which generate obvious and devastating false-negatives.

This same point may also be made in terms of false-positives. To clarify, consider that the slavery and slave-holding maxims incorporate highly specific positions within society and the economy, and this specificity contrasts sharply with the generality of the hospitality maxim. The incorporation of this *kind* of specificity generates false-positives. There are any number of highly specific social and economic positions which cannot be universally willed without contradiction for the simple reason that if everyone were to pursue the same socio-economic position or profession the possibility of achieving this position or profession would be undermined. For example, on this distorted universalization procedure, one could not will to become the CEO of a manufacturing company since, if everyone were CEO with all that that practically entails, then there would be no manufacturing jobs to manage, no raw materials to manufacture from (as all the jobs that extract and supply such material are vacant), etc. Likewise, one could not will to be an astronaut since, if universalized, there would be no general space program (composed of engineers, astrophysicists, and computer programmers) that enable the astronauts to explore the heavens. It is absurd, however, to conclude that we must judge a maxim to become an astronaut, as well as any number of similar maxims, as unworthy!

These consequences are not limited to the slavery and slave-holding examples, but generalize to various maxims. O'Neill illustrates the success of the universalization test when applied to governing principles by considering several putatively impermissible maxims, including the maxims of 'deceit' and 'abrogation of autonomy.' Regarding 'deceit,' O'Neill contends that "a maxim of deceiving others as convenient has as its



universalized counterpart that everyone will deceive as convenient” with the consequence that “there would be no trust or reliance on other’s acts of communication” (O’Neill, 1989, 97). This conclusion, which seems so obvious, does not follow because the maxim does not have sufficient specificity to generate the requisite contradiction. To see this, first consider that the maxim is susceptible to equivocal interpretations of “convenient.” One may mean merely that the agent deceives *when it is easy and requires little effort*. This meaning of “convenience,” however, fails to generate a contradiction as it implies that any rational agent will adopt the ancillary principle of only deceiving the gullible; never the circumspect. Hence, the maxim so construed generates a false-negative.

No doubt, this construal will be rejected: what is meant by “convenient” is that one deceives when such deception *fits one’s plans or purposes*. However, under this construal of “convenient” no contradiction is generated *at all* for two reasons. First, unlike Kant’s original false-promising scenario – which shows how the universalization of a maxim to deceive under very specific conditions establishes an immutable law of deception under those same conditions, thus undermining the agent’s present means – O’Neill’s thin maxim establishes a self-defeating law. Agents, under this law, deceive when it is convenient. But if convenient deception were the universal law, then deception would be ubiquitous, and, consequently, would no longer be convenient (i.e., effectively suit one’s purposes). If no one were to deceive because deception was ineffectual, then “trust or reliance on other’s acts of communication” would still obtain. And yet, as soon as this was sufficiently recognized, deception would then become convenient, and relevant forms of deception would ensue. Thus, instead of generating a clear contradiction, universalizing a thin maxim of convenient deception generates a self-

perpetuating vacillation. It represents a universalization scenario which never reaches a stable conclusion of contradiction or non-contradiction.

Secondly, it is simply not true that people – especially generous ones – will (as a matter of empirical fact) automatically mock any attempt to, say, borrow money under a condition of universal convenient deception. This is because it will not be immediately transparent to the lending agent *why* their associate is asking for money (again, unlike Kant's example, where the dire need for money dictates universal deception, frames the context of deception, and thereby reveals a strict corollary of universal derision whenever one promises repayment). Rather, agents will inquire into the promising agent's *reasons* and *motivations*. If some evidence or explanation can be given for why the current promise is not a matter of mere convenience, and deception is not thereby deduced, people will continue to trust and rely on the borrowing agent's promise as one normally would. Indeed, such a world of universalized convenient deception would not be markedly different from ordinary practices of honesty, trust, or promising where agents, aware of the ubiquity of convenient deception, withhold trust without sufficient warrant *for* trust. In a world of universalized convenient deception it would simply be that some cases of insincere promising would be obvious in certain particular and practical ways, and thus warrant for trust indubitably absent.

Finally, a parallel critique may be made concerning the maxim to 'abrogate one's autonomy' – which O'Neill considers "the most fundamental for a universalization test" (O'Neill, 1989, 96). Simply, O'Neill argues that if everyone were to defer to the judgment of others, which would constitute a universal "elective heteronomy," then, no one could make a decision, and there would be no one to whom anyone could defer.

Notably, the maxim of elective heteronomy does not incorporate any specific socio-economic context. In other words, unlike the slavery maxims which specify a specific socioeconomic position and context, the maxim of elective heteronomy is generic such that the reader is left to his own devices concerning how to conceive of the agent and the conditions under which the agent wills abrogation. Yet the conditionality or unconditionality of the maxim matters when it comes to universalization. Specifically, by “elective heteronomy” one may mean an unconditional abrogation of autonomy; that is, that the agent *always and at all times* abrogates his autonomy. Alternatively, one may mean a conditional abrogation. And yet it is the non-conditional abrogation which is necessary for generating the intended contradiction because if the maxim were to incorporate specific conditions or a context, this conditional abrogation of autonomy could universalize.

This latter point deserves further attention. There are two points to be made here: First, the conditional abrogation of our autonomy is a common feature of numerous actions, a paradigmatic case being joining the military. Within the military (perhaps more so in certain branches or divisions) one's common daily decisions (when to wake, what to eat, what to do and when to do it, how to dress, etc.) are made for the agent. Furthermore, one cannot simply choose to renege on one's decision to enter the military. Importantly, in such cases an agent retains some autonomy; nevertheless, the agent has abrogated autonomy to a significant degree. Furthermore, joining the military is only an extreme case that exemplifies a common experience of abrogating our autonomy by assuming certain positions and especially those involving hierarchical decision-making. Indeed, one might argue that for any decision one abrogates one's ability to choose the

alternative, and thus one's capacity for choice is diminished. Thus, if abrogating one's autonomy to any degree is impermissible, then the universalization procedure will generate untold false-positives.

The second point to be made is that it is not true that no one could make a decision if *conditional* elective heteronomy universalized. First, even when autonomy is abrogated to a degree one still maintains some autonomy so long as the agent can formulate and adopt secondary intentions. That is, autonomy does not concern only first-order decision making but also obtains in our decision-making capacity about subsidiary actions. Accordingly, the consequence of universalization is that not everyone could make every kind or order of decision; and this would require that decision-making of certain kinds and order would need to be distributed according to some social scheme – but this requirement is not contrary to universalization, and thus does not undermine the permissibility of conditional elective heteronomy.

However, by “elective heteronomy” O'Neill more likely means the *total* abrogation of autonomy, and thus the inability for anyone to make a decision if the maxim is universalized. I think this is a point that Hegel could concede while noting the extreme consequences of generating maxims at this level of generality. That is, we can concede that the maxim of abrogating one's autonomy unconditionally cannot be universalized without contradiction. However, this would mean adopting a maxim of elective heteronomy *always and at all times*. But such unconditionality always leads to absurd and contradictory results. Borrowing Allen Wood's example once again, one could not universalize a maxim to unconditionally play tennis – that is, a maxim to play tennis *always and at all times*. Doing so would be practically impossible in light of the lack of

resources both human and material. (If everyone were playing tennis always and at all times, who would make the balls and who would maintain the grounds?) And yet, dedicating one's life to tennis (or some other activity that can only be perfected over a lifetime) is not self-evidently impermissible. Thus, the unconditional formulation of maxims does generate contradictions, but does so at the cost of virtually universal false-positives.

In sum, the dual problem of false-negatives and false-positives has not been circumvented by construing maxims as governing principles. To the contrary, the successful application of the universalization test to thin maxims depends on an inconsistent formulation of the proposed governing principles. When maxims are formulated consistently in terms of the same grade of specificity or generality the universalization procedure is either shown to not generate a contradiction, or to generate both false-positives and false-negatives.

## **B. Multiple Maxims and Moral Subjectivity**

While O'Neill may have developed her view of thin maxims to circumvent the dual problem of false-positives and false-negatives, he is quite explicit about the advantage maxims *qua* governing principles have in avoiding the problem of multiple maxims. The problem of multiple maxims is similar to the problem of gradational indeterminacy – both problems point to the fact that different maxims may be formulated for a given action. The problem of multiple maxims, however, states that a maxim *different in kind* may be formulated for an action, whereas gradational indeterminacy presents the problem as one of *degree*.

O'Neill's own examples are again illustrative. "Notoriously some Nazi war criminals claimed that they were only 'doing their job' or only 'obeying orders' – which are after all not apparently morally unworthy activities" (O'Neill, 1989, 87). O'Neill contrasts these intentions with genocidal intentions which are clearly not universalizable (and which he thinks are representative of these agents' true maxims). Importantly, each intention may serve as the ground for distinct maxims. This poses a serious problem since "it seems easy enough to formulate *some* principle of action for any act, indeed, possibly one that incorporates one of the agent's intentions, which can meet the criterion of any universality test, whatever the act" (87).

O'Neill's assessment of the problem of multiple maxims appears similar to Hegel's moral subjectivity objection:

The immediate consequence of this for itself is that any content one pleases can be subsumed under the good.... Theft, cowardice, murder, etc., as actions... have the immediate determination of being the *satisfaction* of such a will, and hence of being something *positive*. (PR, §140d)

By considering O'Neill's solution to the problem of multiple maxims, however, we can see how he has not appreciated Hegel's deeper objection, and thus has not avoided the overall problem of relying on heteronomous content (in the form, e.g., of a normative specification that proscribes false-promising) in order to successfully apply the universalization test.

O'Neill contends that the problem of multiple maxims does not undermine a universality test for governing principles, because the problem conflates maxims with principles formulated from "surface intentions" (O'Neill, 1989, 87). Of course, the question immediately arises, "Which of the numerous proposed principles are the product of 'surface intentions' and which are fundamental to the act?" O'Neill suggests that we

can adjudicate this question by distinguishing sincere from disingenuous principles.

When considering the ostensibly worthy maxims of Nazi criminals:

The disingenuousness of the claim that such acts were not morally unworthy lies in the fact that these Nazis were not only obeying orders, and indeed that in many cases their specific intentions were ancillary to more fundamental intentions *or principles* that might indeed have revealed moral unworthiness in the agent. (Such fundamental intentions or principles might range from 'I'll do whatever I'm told to so long as it doesn't endanger me' to a fundamental maxim of genocide.) (87)

This rebuttal to the problem of natural actions is unsuccessful for two reasons. First, its success relies on the dubious claim that perpetrators of atrocity have universally adopted, even if they have not expressed, maxims which reveal the "moral unworthiness in the agent." However, we only need one example of a sincere Nazi, or other perpetrator of atrocity, in order to resurrect the problem. Tragically, we do not have good historical reasons to conclude that all perpetrators of gross atrocities are disingenuous when formulating their ostensibly universalizable maxims. While many Nazis, perhaps even all Nazis, committed atrocities grounded in morally unworthy maxims, this is not necessarily the case, and countless other agents have committed and will commit atrocities grounded in sincerely held universalizable maxims like those O'Neill dismisses.

Second, and more fundamentally, O'Neill does not address the origin of maxims, morally worthy or not, and thus fails to appreciate the essential role heteronomous content plays in the function of the universalization test. In other words, O'Neill has passed over the more fundamental problem of heteronomous content Hegel identifies in his moral subjectivity objection.

As I have argued in chapter 1 and the preceding section, Hegel argues for the fundamental subjectivity of maxims, and consequently, the Formula of Universal Law. This argument remains relevant to maxims *qua* governing principles. To reiterate Hegel's

position, there are contending and conflicting descriptions concerning what is essential to any surd action. This ambivalence is simply a consequence of the fact that the 'good' of an action (i.e., the value ascribed to a concrete action) is an aspect determined by each individual agent. Thus, whatever good intention is imputed to a concrete action determines, or rather constitutes, what is essential to the full (i.e., intentional) action. This imputation of what is 'good' in the action is grounded in the agent's arbitrary will, and is thus heteronomous. Said differently, there is no objective ground – available to Kant – for determining what is 'truly' essential to any action, and thus no objective ground for determining which of any number of possible principles of action constitutes the 'real' governing principle. Simply, what I am essentially or fundamentally *doing* is a product of what good I judge to be present in or the aim of my action. For Hegel, this implies that *any action*, no matter how atrocious, may be performed and a maxim may be subjectively imputed to this action which justifies the action *and* shows the moral worth of the agent – as seen with the *Well-intentioned Murderer* discussed above.

This objection contrasts with O'Neill's proposed solution to the problem of multiple maxims: For Hegel, the problem of multiple maxims is not a problem of superficial intentions being erroneously incorporated into a maxim; but rather of agents judging differently what is essential to an action, and thus what ought to be *properly* incorporated into fundamental principles. Thus, the objection O'Neill contends is circumvented by a thin formulation of maxims not only still obtains, it is more fundamental and has more radical consequences. If Hegel is right, it remains "easy enough to formulate *some* principle of action for any act, indeed possibly one that incorporates one of the agent's intentions, which can meet the criterion of any



universality test, whatever the act.” It is the fundamental and not superficial intention of the *Well-intentioned Murderer*, for example, to protect and improve his company.

Additionally, we cannot only formulate ‘some’ principle, but any number of governing principles which can meet the criteria of a universality test. And, most devastatingly, this reveals the inextricability of heteronomy from the universalization test, and the consequences of this heteronomy.

It is important to emphasize this last point. O’Neill has rightly argued that the Formula of Universal Law is not a test designed to determine the *rightness* or *wrongness* of actions, simply because concrete actions are not the target of the universalization test. Rather, the universalization procedure tests the *worth* or *unworthiness* of maxims. Hegel’s moral subjectivity objection may seem to elide this distinction, and so it is critical to see that it does not. Hegel’s position does take seriously the implication that any action whatsoever may be justified without moral condemnation so long as a universalizable maxim is subjectively imputed to that action. And yet, Hegel’s objection has implications specifically for the Kantian concern for moral worth. This is because, though concrete actions do not determine the moral worth of the agent, intentional actions, that is, actions that have been imputed and understood in terms of a maxim, *do*. Thus, Hegel’s objection does not merely conclude that any action may be justified; it also concludes that insofar as we derive the moral worth of an agent from the principle which governs an action, and this principle is determined subjectively, an agent *must* be judged morally worthy (no matter what concrete action he has perpetrated) in terms of this subjectively imputed maxim so long as the maxim is universalizable. Thus, the central feature of the Formula of Universal Law – that we assess the agent’s *own* maxim – has

the extreme implication that we are rationally compelled to judge the perpetrators of monstrous actions (e.g., a war criminal who sincerely claims to be defending his country) to be morally worthy so long as they sincerely act on subjectively determined, universalizable principles of action. For Hegel, this radical subjectivity is inextricable from the Formula of Universal Law.

### C. Ancillary Principles and the Emptiness Objection

I now turn to a unique problem presented by a thin construal of maxims: namely, that ancillary principles also seem to require the application of a universalization test. O'Neill dismisses this worry by pointing out that ancillary principles are not the target of the universalization test. This observation (but not, I argue, the dismissal) is right: the universalization test is intended to test principles of action, and not every particular intention adopted in the realization of an action. If it were, then every agent would be required to test whether they can drive *a blue Honda* to get to a job interview, whether they can live in *the two story house on the corner*, etc. Such tests would be in addition to testing maxims and would mean that Kantian morality requires the testing of multiple maxims. Just as with the problem of gradational indeterminacy with regards to the formulation of a single maxim, testing multiple intentions, at multiple grades of generality and specificity, produces false-positives and false-negatives.

The question of ancillary principles and the problem of formulating a maxim for each intentional action are not new; these were in fact O'Neill's (Neill's) position in *Acting on Principle*, her earlier work on this topic. This inclusive account of formulating maxims was abandoned for reasons we have been considering. Another explanation for

why ancillary principles have not been included as maxims is that the canonical examples of maxims typically incorporate an affected or potentially affected individual or group of individuals in the formulation of the maxim, while many ancillary principles do not. In other words, maxims in the canonical examples are always formulated in terms of actions which involve or incorporate oneself or another. When considering whether to make a false promise, the agent considers whether or not they may make a false promise *to another agent*. When considering if they may keep a deposit for which there is no proof, they are asking about a specific deposit that is the property of an individual person. This consideration of another is inherent to these cases: proposing to keep private property presupposes the private party, an agent or entity, entitled to the property; false-promising is a communicative act that requires at least one other agent. In cases of developing talents and suicide, the person affected is the agent himself. Thus, in all of these cases, the agent's maxim is either self-directed or other-directed in such a way that an agent is incorporated into the maxim as the intentional object.

This intentional structure constitutes an intuitive threshold for when principles of action require testing through universalization. For any proposed principle of action, if the principle of action implicitly or explicitly incorporates an agent (whether the acting agent himself or another) as an intentional object, then the principle of action requires the application of a universalization test. Thus, if my maxim is to *use a car without permission* it must be tested; if the maxim is merely to *drive a car* it need not be tested. If this condition were true, it would explain why a universalization test is not and need not be applied to ancillary principles, such as *stirring coffee* or taking a *Honda sedan*: ancillary principles of this kind simply do not include morally germane content because

they do not include agent-centric intentions.

Construing maxims thinly establishes a different threshold for when principles of action must be tested, and thus brings to light the problem of allowing ancillary principles – that is, of not having an inclusive account of maxim testing. Although thin governing principles, like thick maxims, are intentional about agents, under a thin interpretation of maxims not *all* principles of action which are intentional about agents are subject to a universalization test. Consequently, on a thin interpretation of maxims, some ancillary principles which incorporate specific intentions about agents are not subject to the universalization test. This threshold introduces a question concerning the permissibility of ancillary principles. Simply put, if my ancillary principles are no longer about objects or beings which are not of moral concern, but about agents who require moral consideration, then the ancillary principles in service to my governing principles also require moral evaluation – that is, they must be subjected to a universalization test.

O'Neill's example of hospitality is again illustrative. In this example, the maxim under consideration is “make a visitor feel welcome.” Ancillary principles considered include offering coffee and the various intentional actions – such as selecting a mug – which serve to realize this action. All such actions are benign. But it is obvious that this need not be the case: there are countless ways in which a visitor may be made to feel welcome which involve subordinate actions and ancillary principles that are intentional about agents and which are not benign. To begin with a trivial example, consider gossiping. A host wishes to ‘break the ice’ with some engaging conversation and so decides to share some gossip about a mutual acquaintance. The ancillary principle of this subordinate action would be “report unconfirmed yet scandalous or embarrassing

information about our mutual acquaintance.” Can this principle of action be universalized? The answer is uncertain. However for our purposes the crucial point is that the questionable principle of action is not even a candidate for universalization. Because the gossiping principle of action is ancillary to the agent’s maxim it is (on a thin interpretation of maxims) simply not the target of a universalization test.

I have chosen the mundane example of gossiping as a heuristic, but it is evident that the general problem has disastrous consequences for Kantian ethics if we assume an unaltered thin interpretation of maxims. Consider briefly the Eskimo practice of wife-exchange raised by James Rachels (1999) and nearly ubiquitous in introductory ethics courses. In outline, it was a customary form of hospitality for an Eskimo man to offer his own wife to a guest of a certain connection or under certain conditions.<sup>15</sup> On a thin interpretation of maxims such an act would be subordinate to the overall intention to be hospitable, and the correlative ancillary principle would not be subject to universalization. A thin interpretation of maxims, then, sequesters questionable and even obviously impermissible maxims from the purview of the universalization test.

These brief examples are simply representative of a fundamental challenge to thin maxims: on a thin interpretation of maxims any principle of action whatsoever, so long as it is genuinely ancillary to a universalizable governing principle, is permitted.

The solution to the problem of ancillary principles is obvious: test ancillary principles that exhibit an intentional structure incorporative of any agent. However, this means testing thick principles of action – that is, principles of action which include

---

<sup>15</sup> As noted in the text, I am adopting the claims made by James Rachels in his *Elements of Moral Philosophy* (1999, 15-29), and which have become canonical examples of cultural relativism. For anthropological support of these claims, see Rubel (1961), available: <http://www.uaf.edu/files/apua/Rubel1961.pdf>.

specific intentions – even though maxims per se are still construed thinly. Returning to the example of hospitality, this would mean that an agent must test their maxim of “making a new visitor feel welcome,” but must also test any ancillary principle in service of their maxim which exhibits the relevant intentional structure, such as “share gossip about a mutual acquaintance” or “tell a flattering falsehood to my guest” and perhaps even “offer coffee to my guest.” Importantly, since such ancillary principles include specific intentions, they are susceptible to the same emptiness critique leveled against thick maxims: Whether such ancillary principles pass or fail the universalization test will depend on the particular normative specification incorporated into the agent’s will and maxim. Moreover, if the ancillary principle fails the universalization test, it is because the normative specification, posited prior to the test, already includes the relevant proscription. As argued extensively in the preceding sections as well as chapter 1, this condition for success renders the result of the test tautologous and the test itself superfluous.

In short, then, the problem of ancillary principle presents the Kantian with a dilemma: maintain a thin construal of maxims allowing for the adoption of impermissible ancillary principles; or, require ancillary principles to pass the universalization test with the implication that this will generate only tautological results and have the consequence of permitting impermissible ancillary principles under the right normative conditions.

### III. CONCLUSION

In this chapter I have argued for the first stage of extending Hegel’s emptiness objection by demonstrating the effectiveness and relevance of the alternative interpretation

developed in chapter 1 for the Formula of Universal Law generally. Specifically, this chapter applies Hegel's emptiness objection to additional canonical cases illustrative of the universalization test and considers whether Hegel's objection remains relevant for different interpretations of maxims. Regarding the first goal of extending Hegel's objection to other canonical cases, Hegel's arguments were effectively adapted to the false-promising scenario, and arguments paralleling those developed in chapter 1 demonstrated how the success of the universalization procedure requires that a normative specification inclusive of the proscription against false-promising must be posited prior to the implementation of the universalization procedure. I then turned to a consideration of the relevance of Hegel's critique for Onora O'Neill's interpretation of maxims as governing principles. I argued that with regards to emptiness objection and negative duties, Hegel's emptiness charge cannot be directly applied to maxims *qua* governing principles. However, maxims so construed face a number of problems, including gradational indeterminacy, multiple maxims and moral subjectivity, as well as the status of ancillary principles vis-à-vis the universalization procedure, and in light of these challenges I argued for the relevance of Hegel's critique. It now remains to be seen whether Hegel's critique can be extended to positive duties as well as other formulations of the Categorical Imperative – a problem I take up in chapters 3, 4 and 5.

### **Chapter 3:**

#### **Does the Formula of Universal Law Show Too Much?**

The previous chapters have focused exclusively on the Formula of Universal Law and negative duties. In this chapter I remain focused on the Formula of Universal Law but consider what the formalism of the universalization procedure means for maxims of positive duty. Hegel argues, as part of his emptiness critique, that the formalism of the universalization procedure may rule innocuous and even obligatory maxims impermissible. This aspect of Hegel's emptiness objection is sometimes dismissed; I argue that, on closer examination, it is a tenable and important objection.

In Kant's own lifetime the Formula of Universal Law was challenged as an effective means of determining duty insofar as the universalization procedure would, under the right circumstances, restrict our actions in ways contrary to what is intuitively permissible or even obligatory. This challenge was most famously and effectively made by Benjamin Constant in 1796 with the introduction of his example of a "murderer at the door" and Kant's subsequent, apparently rigoristic, response in *On a supposed right to lie from philanthropy*. In his *Natural Law* treatise, Hegel presents a similar objection presented as a corollary to his emptiness objection; he argues that the abstract emptiness of the universalization procedure will sometimes falsely generate a contradiction and thereby proscribe the performance of positive duties. This chapter explores the effectiveness of each objection, and defends both against proposed solutions in recent Kantian literature.



With regards to Hegel's objection, the preceding chapters have presented an alternative interpretation of Hegel's emptiness objection and considered the extent and relevance of this alternative interpretation for different interpretations of the Formula of Universal Law. I have argued that Hegel's emptiness objection must be understood in terms of a constellation of arguments, and is thus best understood as an emptiness *critique*. However, even with an appreciation of Hegel's supplementary arguments, the full extent of Hegel's position has not yet been presented. This is because Hegel also argues for a further consequence of the emptiness of the Formula of Universal Law: namely, that *without* assumed, particular content the Formula of Universal Law *shows too much*.<sup>1</sup> This objection – when it is addressed at all – is generally considered independently of the emptiness objection. I think this is a mistake, and one aim of this chapter is to show how the two objections are related. For the sake of clarity, however, I will refer to this corollary objection as the “excessiveness objection.”

The excessiveness objection may be considered well-trodden territory; as we will see, it is essentially an iteration of the false-positives objection considered in chapters 1 and 2, which is a common concern within the Kantian secondary literature. However, as presented by Hegel, the excessiveness objection is importantly unique. First, false-positives are sometimes presented as a problem related to the generation of *negative* duties; that is, false-positives are generated that restrict maxims and correlative actions that are in fact permissible. In contrast, Hegel's objection is presented exclusively as a problem for maxims that satisfy our imperfect positive duties, and thus a problem for actions that are *obligatory*; one example is the maxim of generosity in the service of our

---

<sup>1</sup> I borrow this phrase from Korsgaard (1996, 86-87).

duty of beneficence.<sup>2</sup> Second, false-positives are considered a product of maxim gradation; that is, a false-positive is generated when a maxim is formulated with a particular (usually high) degree of specificity. Thus, while a solution to the problem of false-positives has not been successfully developed (i.e., at what degree of specificity a maxim should be formulated has not been settled), the solution to the problem of false-positives is clear in theory: simply formulate maxims at the right degree of specificity. Again, the excessiveness objection contrasts with this standard presentation of the problem: Hegel's examples of maxims that wrongly generate a contradiction are formulated either with a low degree of specificity, or with a degree of specificity comparable to Kant's own examples.

The uniqueness of the excessiveness objection provides sufficient reason to give it independent consideration. This reason is amplified by the abbreviated attention the excessiveness objection receives in the secondary literature. Most importantly, despite the fact that Hegel explicitly states that the generation of excessive contradictions is a problem for an "infinite" number of maxims, and despite the fact that false-positives are a known problem for the universalization procedure; the excessiveness objection is framed and rebutted in terms of a single example and not in terms of a generalizable objection. This chapter, then, aims to rectify the shortcomings within the secondary literature by both interpreting the excessiveness objection in light of Hegel's emptiness critique, as well as explicating Hegel's objection in terms of a generalizable, formal structure which

---

<sup>2</sup> I have already presented this problem, in brief, in chapter 2 (pgs. 82ff). In response to O'Neill's suggestion that both a maxim of slavery and a maxim of slave owning are impermissible, I argued that other maxims that incorporate similarly specific social positions (such as a CEO or an astronaut) are equally impermissible. This chapter furnishes the opportunity to consider this problem, and proposed solutions, in more detail.

provides a template that may be applied to any number of maxims.

In contrast to the paucity of commentary on Hegel's excessiveness objection, Constant's "murderer at the door" example is one of the most discussed objections to Kant's moral theory, and undergirds a general attribution of rigorism. I consider two very different approaches furnished to rebut this rigorism. Christine Korsgaard (1996) argues that the Formula of Universal Law can accommodate the murderer at the door example insofar as it represents a case where evil is already present in the form of deception, and such antecedent deception changes the context in such a way as to allow for universalization of a lying maxim without contradiction. Recently, in a very different approach, Allen Wood (2008) has presented an exegetical rebuttal to the attribution of rigorism by arguing, first, that Kant's specific response to the 'murderer at the door' example must be understood contextually as a proscription against false *declarations*, which within Kant's moral theory are statements made within a context (such as an oath of office or legal testimony) where one has warrant to believe the speaker and the speaker is liable for the truthfulness of the statement; and second, that Kant's general rigorism can be attributed to pedagogic hyberbole. Properly interpreted, the apparent rigorism of Kant's response to Constant disappears and his general rigorism has a new, legitimate purpose. I respond to both Korsgaard's and Wood's rebuttals and argue that each fails to successfully resolve the crucial issue at work in Constant's challenge.

The chapter thus proceeds as follows: In Section I I first reconsider the excessiveness objection, explicating it as a corollary to Hegel's general emptiness critique. Next, I provide a more detailed explication Hegel's argument with particular emphasis on his different examples and then contrast this interpretation with competing

interpretations in the secondary literature; finally, I evaluate the effectiveness of Hegel's objection in light of different interpretations of the Formula of Universal Law. In Section II I present Constant's famous example as well a critique of Korsgaard's and Wood's proposed solutions. I conclude that (1) while some interpretations of the Formula of Universal Law (*viz.*, the Logical Contradiction Interpretation) have limited success in rebutting the excessiveness objection, this success does not generalize to other examples, nor is it shared by other interpretations (*viz.*, the Practical Contradiction Interpretation); and (2) no interpretation (thus far) can avoid the unacceptable rigorism of the Formula of Universal Law.

### I. HEGEL'S EXCESSIVENESS OBJECTION

Concerning the independence of the excessiveness objection, it is perhaps unsurprising that Hegel's arguments concerning the emptiness of the Formula of Universal Law with regard to positive duties differ from those concerning negative duties. Negative duties determine what is forbidden, or the limit of permissible action. Positive duties determine what is required, or the limit of our permissible *inaction*. The consequences of emptiness correlate to this distinction. In the case of negative duties, Hegel contends that without incorporating illicit content the universalization procedure will fail to limit impermissible maxims: false promising will only be judged impermissible given the *right* normative specification; namely, one that already includes a proscription against false-promising to escape financial need. Without this normative specification, or with a different normative specification, the surd act of telling a falsehood, even for one's financial self-interest, may be permitted. In contrast, Hegel argues that in the case of positive duties, maxims which

ought to be permitted – specifically, maxims of imperfect duty – will sometimes generate a contradiction and be judged impermissible. Thus, instead of failing to limit impermissible maxims (e.g., maxims of stealing or false-promising); the universalization procedure proscribes *obligatory* maxims of imperfect duty. In this way, Hegel argues that not only is the Formula of Universal Law insufficiently content-full for determining, with concreteness, what is *moral*; this same emptiness also produces *immoral* conclusions (see *NL*, 80).

In order to contrast the two objections it may be helpful to briefly review the argument developed in chapters 1 and 2. Hegel argues that the universalization procedure elevates the “specific thing” – the proposed maxim – to the universal, and thereby tests whether a specific content of the agent’s will and maxim may be made universal law; but this elevation of the specific and subsequent annihilation requires a normative specification that already includes the relevant proscription such that the universalization procedure produces only tautologies: if property understood as a certain normative practice is assumed, then one cannot will a maxim which violates the norms of that normative practice. But this only shows that a maxim is permissible or impermissible relative to a presupposed normative practice. Moreover, the universalization procedure functions equally well in any ethical life or in terms of any normative specification, including an ethical life that incorporates as normatively permissible the action proscribed. Lastly, the normative specification required for generating the desired contradiction (which would, e.g., require honesty or a particular respect for property) must be determined prior to the universalization procedure itself. Thus, the universalization procedure is, in itself, *empty* in that it is nothing more than a formal

relation that may be applied to any substantive normative context, and thus, may prescribe or proscribe any particular *concrete* principle of action whatsoever.

Despite the apparent differences between the emptiness and excessiveness objections, each are grounded in the emptiness of the Formula of Universal Law, and are best understood as corollaries of the general emptiness charge. This claim is supported both conceptually and exegetically. Regarding the latter justification: in the passages of *Natural Law* immediately following this contention, Hegel argues that this same formal emptiness faces another, even more direct problem: “This annihilation of the specific, through the adoption into infinity and universality, is indeed an immediate difficulty for practical legislation” (*NL*, 79). By “*this* annihilation of the specific” Hegel is referring back to the universalization procedure and the determination of a particular proposed maxim as impermissible. In addition to this direct reference to the same procedure, both emptiness and excessiveness are products of the same “adoption into infinity and universality.” This is to say that both are the consequence of same formal emptiness of the Formula of Universal Law. The difference lies in the way this emptiness functions to show too much in the case of positive duty. Specifically, the generation of a contradiction in the case of maxims corresponding to negative duties and maxims corresponding to positive duties has an inverse relationship to content and the Formula of Universal Law: in the former case, it is the presence of content that produces a contradiction (a “true” positive); in the latter case, it is the absence of content that produces a contradiction (though a false-positive). In the case of maxims corresponding to negative duties, when the universalization procedure is utilized to show the impermissibility of a maxim, the generated contradiction is made possible by restricting the proposed maxim to a posited

normative specification; it is the supplied, illicit content that generates the specific contradiction. In the case of maxims corresponding with positive duties, a contradiction is generated because of the unrestricted universality of the Formula of Universal Law; the pure universality of the Formula of Universal Law, empty of specific social, historical or discursive norms, generates a contradiction of will from select maxims of positive duty. In this way, the formal emptiness of the Formula of Universal Law both cannot rule out impermissible maxims, but also cannot ensure the inclusion of required maxims. It is, moreover, in this sense that we should read Hegel as arguing that the universalization produces immoral results: false-positives that are contrary to our positive imperfect duties.

This charge of excessiveness is best explicated by considering Hegel's own examples, which I turn to presently. I first consider Hegel's example of beneficence – the maxim to “Help the poor” – which is most often discussed in the secondary literature. However, I use this example to establish a template for the objection, and then apply this template to Hegel's other example as well as other putatively permissible maxims – both original and drawn from the secondary literature.

### **A. Superseding a Supersession: Clarifying the Structure of Hegel's Objection**

As already noted, Hegel's excessiveness objection is often interpreted in terms of a single example. I will also begin by explicating Hegel's position in terms of his sole developed example, but then turn to the formal structure of the excessiveness objection which this example clarifies. Turning to Hegel's second though undeveloped example further provides an opportunity to rectify what I perceive to be a misinterpretation of Hegel,

which in turn grounds an ineffectual Kantian rebuttal.

Turning to the text, Hegel asks the reader to consider the maxim, "Help the poor." He states that this maxim "expresses the supersession of the specific thing, poverty" (NL, 80), by which Hegel means that the maxim of generosity (which I take to be in the service of duty of beneficence) has the intent of eliminating poverty. (Whether Hegel here means poverty *simpliciter* or a particular instance of poverty is not made clear; this is a distinction that has some import for responses to Hegel's objection considered below.) Hegel continues, arguing that if this maxim is universalized it results in a contradiction. This contradiction is brought about because if helping the poor were a universal law then *either* there would not be any poor, or everyone in their attempt to help the poor would themselves become poor. In either case helping the poor would be practically impossible.<sup>3</sup> In this way, the maxim which itself expresses a supersession of poverty is superseded by universal legislation, and this universal legislation in superseding the original maxim cancels the necessary conditions (*viz.*, poverty *or* affluence) for its implementation and thus denies the permissibility and dutifulness of the maxim.

Crucial to the understanding of both Hegel's position and rebuttals to his position is the fact that this example is not an independent objection but represents a general problem whereby a maxim that ought to be permissible, and would in fact satisfy a positive duty, will be ruled impermissible by the universalization procedure. Hegel presents the problem formally in a particularly difficult passage:

For if the specific thing is such that in itself it expresses the supersession of something specific, then, by the elevation of the supersession to universality or to the state of having been superseded, not only the specific thing which is to be

---

<sup>3</sup> Hegel goes on to point out that if poverty were to remain it would remain for the purpose of making possible the maxim to help the poor, but that "this maintenance of poverty forthwith means that the duty is not fulfilled" (NL, 80).



superseded, but the superseding itself, is cancelled. Thus a maxim referring to such a specific thing, which cancels itself when it is universalized, would not be capable of being the principle of a universal legislation, and so would be immoral. (NL, 80)

It is not clear whether Hegel intends this analysis to apply to all maxims, though in order to illustrate the objection we can see how this structure of a double supersession and cancellation may be used to describe other canonical scenarios. For example, returning to the false-promising scenario, the deceitful agent adopts a maxim with the intent of escaping from and thus replacing a current financial difficulty. Employing the maxim would result in the acquisition of money, and the state of affairs of *having money* takes the place of (supersedes) a specific financial difficulty. Additionally, when the maxim is universalized the universal legislation supersedes the individual's maxim. That is, in the world of the universalized maxim universal legislation takes the place of an individual maxim. This latter supersession, however, may lead to the cancellation (to use Hegel's terminology) of the necessary conditions for acting on the specific maxim: in the case of false-promising, the specific maxim led to a contradiction because the ground for promising (i.e., trust) was annihilated. Likewise, in the case of some maxims of positive duty (e.g., the duty of beneficence) a contradiction is generated because the grounds for acting on the maxim "Help the poor" are canceled. In other words, the conjunction of affluence and some poverty are the necessary conditions for helping the poor and one of these conditions is annihilated under a universal law of generosity to the poor. As in the case of all maxims, if the necessary conditions for acting<sup>4</sup> on a maxim are canceled by the

---

<sup>4</sup> Here I leave the characterization of impermissible maxims in terms of the Practical Contradiction Interpretation. As I will discuss below, other interpretations – namely, the Logical Contradiction Interpretation – provide an ostensible way of avoiding this contradiction and thus Hegel's objection. It is also worth noting, first, that this analysis of Hegel's objection reinforces the contention in chapter 1 that Hegel understood how the universalization procedure works, with regards to content being

universalization of the maxim, and this generates a contradiction, then the proposed maxim is impermissible. Thus, the excessiveness of the universalization procedure is revealed: at least some maxims of positive duty may in fact be formally contrary to duty.

Hegel claims that this structure – the supersession of a supersession – applies to an “infinite number of other maxims” (*NL*, 80). However, Hegel only provides one additional example. Considering this additional example has special import, as it not only aids in further clarifying Hegel’s position, but also serves as a possible means of rectifying a misinterpretation of Hegel’s argument which further undergirds a certain kind of Kantian response.

The maxim Hegel briefly proposes is that of a patriot:

So the maxim of honorably defending ones country against its enemies...is self-cancelling as soon as it is thought as a principle of universal legislation; for when so universalized, for example, the specification of country, enemies, and defense is cancelled. (*NL*, 80)

Taken independently of the general critique, one might interpret this passage to mean that the patriotic maxim inherently cannot be thought as universal legislation because as soon as it is the specificity of country and enemy is lost. That is, from a universal standpoint there is no specific country identified in the maxim, but rather a generic loyalty to a generic state. But this is simply not how patriotism works. And indeed, this passage has been read to mean that patriotism concerns one’s own, unique country, and thus there is a “contradiction” in universalizing patriotism because the maxim, as formulated, renders patriotism generic – that is, it characterizes patriotism as loyalty to any country

---

provided by the maxim; and, second, not only does Hegel’s critique remains relevant for internal contradictions of will, or the Practical Contradiction Interpretation, but, further, Hegel’s interpretation of Kant appears to be a variation of the Practical Contradiction Interpretation proposed by Korsgaard over 150 years later.

whatsoever. Representative of this view<sup>5</sup> is H. B. Acton's (1975) commentary on this passage in the introduction to *Natural Law*:

Hegel's point, I think, is that each country is unique and that the individual's loyalty is to it rather than to some principle about patriotism generally....By patriotism the individual is attached to those particular men who form this particular society rather than to all patriots everywhere. (27)

This interpretation, in turn, enables a certain Kantian reply; namely, that "we can and should always ask whether the action in these circumstances could be carried out by everyone who found himself in them.... [One's country] might, for example, be thoroughly corrupt, and its enemies might have justice on their side" (27). That is, the maxim as formulated by Hegel is already unacceptable insofar as a person must, antecedently, ask whether patriotism is permissible (universalizable) in light of the policies and actions of the country or state in question. One must first ask whether patriotism for one's country is permissible before asking whether honorably defending one's country is permissible. Thus, on an interpretation that takes Hegel to mean merely that patriotism is and cannot be generic, this fundamental Kantian question has not been answered and Hegel's objection appears misplaced.

While this interpretation may accord with Hegel's general view<sup>6</sup> of patriotism, and while it may also be true that there is an inherent contradiction in a generic, stateless

---

<sup>5</sup> See also Franco (1999, 60).

<sup>6</sup> Hegel's view of patriotism and critique of Kant's cosmopolitanism is beyond the scope of this paper. For a thorough account and sympathetic defense of Hegel's critique of Kant's cosmopolitanism, see Fine (2003). While Acton's brief characterization does accord with Hegel's view, it gives a superficial presentation. Hegel's full view of patriotism is not grounded in the mere uniqueness of each country or the natural loyalty among comrades. Rather, patriotism, according to Hegel, is a "political disposition" founded in "the consciousness that my substantial and particular interest is preserved and contained in the interest of the other (in this case the state), and in the latter's relation to me as an individual" (*PR*, §268). This trust results in the state ceasing to be in the 'other', and thereby contributes to the realization of the agent's genuine and rational freedom.

patriotism; the interpretation nonetheless fails to conform to the structure of the specific objection presented in the *Natural Law* treatise and thus fails to appreciate the relevance of Hegel's objection. Any genuine interpretation of the excessiveness objection must conform to the formal structure Hegel has provided, and thus must disclose how there is a supersession of a supersession, and how the agent's proposed maxim is ruled impermissible in virtue of the fact that the necessary conditions for acting on the maxim are canceled through universalization.

We may meet this requirement by elaborating on Hegel's brief example. To do this we must first establish the specific thing the agent's maxim intends to supersede. Hegel does not provide the intention of the maxim, though for the sake of illustration we may say that a patriot seeks the security and preservation of his country; and thus, conversely, the patriot's maxim intends to supersede any real or perceived insecurity and danger to his country. Second, we must elucidate the implications of universalization. In Hegel's terminology, the specific patriotic maxim is superseded by a universal law of honorably defending one's country; this universalization of the maxim implies that there is and has been the honorable defense of *everyone's* country by *every* individual (in the same way that there is and has been universal generosity to the poor to the extent that poverty is abolished or affluence is eliminated). The final step is to ask what necessary conditions are canceled by such universal patriotism.

Interestingly, a variety of answers may be given to this question. Robert Stern (2012) has suggested a straightforward answer: the universal *defense* of one's country implies that there is no enemy *attacking* any country, and thus no need or possibility of defense. In order to defend one's country there must be something to defend it from; this

condition implies the presence of an agitator, but if there is no agitator there can be no actual defense. Thus, the necessary condition for acting on the patriotic maxim is canceled. I do not think this analysis of the consequences of universalization is necessarily wrong; however, it does not precisely follow Hegel's explication. Hegel's view is that the "specification of country, enemies, and defense is cancelled" and it is not clear how, on Stern's view of universal defense without offense, this cancellation is brought about. How is the specification of one's *country*, for example, canceled by universal defense of one's country?<sup>7</sup>

A second possibility is that the universalization of a maxim of honorable defense results in universal war – at least among those for whom there is a real or perceived threat. If we again follow the example of beneficence as well as the abstract structure of Hegel's objection, we can initially identify two concrete consequences of universal war that undermine the possibility of honorably defending one's country: (1) victory, in which case one's enemies *qua* national enemies are annihilated; and (2) defeat, in which case one's country, at least when understood as a state capable of mounting a defense, is annihilated. In both cases the possibility for defense is also canceled. Thus the agent, on universalizing his maxim, wills a contradiction in that he wills to honorably defend his country against his enemies and wills *either* that he has no country to defend or that he has no enemies to defend his country against. A difficulty with this view is that it leaves open a third possibility: perpetual stalemate. This option makes the patriotic maxim practically possible and, thus, non-contradictory.

---

<sup>7</sup> It is also worth noting that this interpretation would be vulnerable to the response, furnished by Korsgaard, from the standpoint of the Logical Contradiction Interpretation. That is, on this interpretation the patriot would simply "have nothing to do"; but having nothing to do is not contradictory on this interpretation.

A third and final answer is more successful both in terms of generating a contradiction and following Hegel's explication. As we have seen, Hegel states that it is the *specification* of one's country, enemies, and defense which is cancelled. This detail is important. A country (and one's enemies) specified is not a mere collection of people; it is, rather, a precise identification of a people and culture with particular institutions (economic, governmental, and social) as well as particular relations between these different elements. Likewise, the specification of "defense" is not the mere act of defense; rather, in this context it is the particular defense *of* one's country as specified. If the honorable defense of one's country were universalized such that *everyone* were to defend their country, then one's own country, as it is currently specified, would be transformed and, correlatively, the possibility of defense undermined. That is, to be concrete, suppose a country is organized such that some people maintain the family (mothers, fathers, and grandparents) while others produce good and generate wealth (laborers, craftsmen, and entrepreneurs) and still others govern (legislators, executives, judges), but then a universal maxim of patriotic defense is adopted such that everyone joins the military and society – in its entirety – is organized militarily; such an adoption would fundamentally change the nature and character of one's country. It is not that such a country is inconceivable; rather, it is that such a country would be unrecognizable. The same transformation would be entailed for one's enemies. Moreover, the defense of one's country would also no longer be possible insofar as one's country *as specified* would no longer exist. In this way, the necessary conditions, which are the specification of one's country, enemy, and defense, are cancelled and the maxim becomes practically impossible.

One might respond that the generation of a contradiction is contingent: we can conceive of an agent who is the citizen of a country already organized around universal military service. This is, I think, true; but it is a moot point. Hegel's objective is not to show that the patriotic maxim is always non-universalizable; but, rather, that the maxim is simultaneously *sometimes* intuitively permissible and non-universalizable. That is, sometimes the patriotic maxim is one that an agent should be able to follow and perhaps has an obligation to follow, and this patriotism is proscribed by the Formula of Universal Law.

Though I clearly prefer the third interpretation, with each of the above interpretations we can now see how Acton's Kantian rebuttal is misplaced: Hegel's objection does not fail to see that we must ask whether our patriotic impulse is consistent with duty and justice; rather, as in the case of the maxim of beneficence, Hegel's objection targets the formal permissibility of a patriotic maxim even when it is consistent with duty and justice. In other words, the aim of the excessiveness objection is to show how our actual positive duty – such as a duty of patriotism or beneficence – conflicts with our formal duty as determined by the Formula of Universal Law; that is, how the universalization procedure, when applied to certain maxims of positive duty, generates false-positives. In this way, to leave unanswered the fundamental Kantian question regarding the justice of one's patriotism would be contrary to the objective of the excessiveness objection which is to show that even when one *ought* to be patriotic (i.e., when the antecedent question of justice has been answered in the affirmative) the universalization procedure rules such patriotism impermissible. Of course we may still ask, and must ask, whether actually fighting in defense of one's country is consistent with

justice, and an answer to this question may even supersede an agent's patriotic motivations. However, in light of Hegel's objection this question would be answered prior to proposing the current maxim of honorably defending one's country. And thus, this obligation clearly does not mean that Hegel has not addressed whether agents in the same circumstance may adopt the same maxim as this is precisely the context in which Hegel raises his objection.

Thus far I have only considered maxims formulated and presented by Hegel; to support Hegel's claim concerning the extensiveness of the excessiveness objection, it may be helpful to apply the formal structure to other maxims. Accordingly, we may briefly consider two maxims of virtue. First, we may again consider the case of an individual who pursues a specific vocation, such as a CEO or astronaut, which was briefly touched on in chapter 2. In this current context, we may stipulate that the agent chooses a vocation in order to satisfy their duty of self-perfection. Second, we may consider an agent who, *sympathizing* with the condition of orphans or neglected children and motivated by a duty of *beneficence*, proposes the maxim "to ensure the happiness of neglected children, I will adopt a child." Both maxims I take to be obviously permissible and plausibly dutiful. Both maxims also produce a false-positive when the universalization procedure is applied.

To see this we must again follow Hegel's formal structure by first identifying the "specific thing" the maxim intends to supersede. In the case of pursuing a vocation, the maxim intends to supersede the current lack of self-development; however, in this case the agent has set a specific end in order to meet this supersession. We may suppose that the agent pursues all of the steps necessary to develop the skills and talents necessary to



become a CEO, and in Kant's language "make him a human being useful for all sorts of purposes" (*G*, 4:423). Accordingly, the agent gets an appropriate undergraduate education, perhaps an MBA, seeks the appropriate work experience, etc. The next step is the supersession of this initial supersession: universalization of the maxim supersedes the individual's specific maxim since it is now a universal law to pursue these same talents and the same means to cultivating these talents. Next one must identify how the necessary conditions of this particular mode of self-perfection are undermined by universalization. Unlike both the beneficence and patriotic maxims, where there were two possible consequences of universalization, in the case of self-perfection we can identify one consequence which undermines one necessary condition of becoming a CEO: namely, that if everyone were to universally and unconditionally adopt the particular maxim of self-perfection (to become a CEO), then everyone would necessarily be a CEO or attempting to be a CEO; and if this were the case, then there would be nothing to be a CEO *of*. Without the workers, producers, designers, financiers, etc. that make the existence and functioning of a business possible, there could be no business and, thus, no CEO. The universal adoption of the maxim to become a CEO annihilates *the necessary conditions for the possibility of being* a CEO – namely, the existence of the diverse constituent elements of a business and market. In this way the excessiveness of the universalization procedure is again revealed: what is, to many, a permissible activity is formally contrary to duty.

The same analysis may be applied to the maxim to adopt a child. The specific thing to be annihilated is the suffering and neglect experienced by an orphan or other child in need. The individual maxim is superseded through universalization when the

principle “adopt a child” is made a universal law. However, because adopting a child is now a universal feature of the agent's society, the necessary conditions for adopting a child disappear – there are no longer any orphans or neglected children to adopt. Thus, again, following the form of Hegel's objection, an intuitively permissible and plausibly dutiful maxim would be ruled impermissible.

The foregoing examples have only clarified what I take to be the structure and applicability of Hegel's excessiveness objection: when the universalization procedure is applied to maxims of positive duties false-positives are consistently generated because through the supersession of a supersession the necessary conditions for acting on a maxim are undermined. Now that Hegel's excessiveness objection has been clarified, I turn to a critical evaluation of his position, including the consideration of relevant rebuttals.

## **B. Objections**

Objections to the foregoing interpretation take a variety of forms, starting with the success of Hegel's examples as well as the application of the excessiveness objection to other maxims of virtue. Of particular importance is a rebuttal concerning the appropriate formulation and integration of conditionality into certain maxims. Once again Korsgaard's Practical Contradiction Interpretation figures centrally. This is because, first, Hegel's own objection seems to be framed in terms of practical contradiction – helping the poor becomes impossible for the agent, but not inconceivable. Second, in the context of defending the Practical Contradiction Interpretation Korsgaard raises what I take to be the most robust rebuttal to Hegel's excessiveness objection. I consider this rebuttal last,

and argue that, while it presents a successful solution to the beneficence maxim, this solution does not generalize to other examples and, further, entails unacceptable consequences for the universalization procedure generally. Before turning to objection based in the Practical Contradiction Interpretation, I first consider some general objections as well as a rebuttal based in the Logical Contradiction Interpretation.

### *1. Do Hegel's examples succeed?*

Even if Hegel's objection in the abstract is sound, one may still question whether his specific examples succeed; that is, whether his analysis of the proposed maxims and consequences of universalization is correct. If Hegel's analysis of these examples fails, one may further ask whether there are any relevant cases of positive duty that actually problematize the Formula of Universal Law.

Objections along these lines are various. First, one might argue that a maxim to help the poor does not have the consequences Hegel thinks simply because a maxim to help the poor does not equate to a principle to help the poor *without limit*; one may simply have the maxim of "giving to the poor." In response, it is not Hegel's claim, nor need it be to problematize the Formula of Universal Law, that no maxim of generosity is universalizable. Rather, all Hegel needs to claim is that ordinary maxims – that is, the kind of maxims that might be formulated by the "untutored" – will sometimes falsely generate a contradiction. Thus, if Hegel's formulation of the maxim, which includes the aim of abolishing poverty, is reasonable, then this consequence does follow. Furthermore, other ordinary maxims such as the maxim to "adopt a child" are more restricted but similarly lead to false-positives.

Second, one might object that the consequences of universalization do not follow, since (due to incompetence, disposition, tragedy, or disaster) there will always be poor. This objection is ineffective for two reasons: First, it reveals the social, historical, and economic contingency of the universalization procedure. In other words, whether or not the maxim to help the poor (or to adopt orphans) is universalizable depends, not on the exceptional character of the maxim itself, but on the character of the society and period of history. Moreover, a similar contextual contingency may be shown for other standard cases.<sup>8</sup> Second, this objection mistakenly takes technical and possible exceptions to the consequences of universalization to undermine the practical consequences of universal legislation. However, this is of no consequence for the agent attempting to will to help the poor, for his hypothetical imperative is still frustrated by the universalization of his maxim – that is, it remains true that, even if there are some poor somewhere, there is nonetheless universal help for the poor and, as such, the agent cannot follow through on his particular maxim in his present circumstances. Said differently, the fact that there will always be poor does not equate to an extant opportunity – and especially a context specific opportunity – to help the poor.

Next, consider the maxim to become a CEO. Some might rightly point out that there is no duty to pursue any particular vocation or activity, and that it is this added specificity

---

<sup>8</sup> In chapter 2 I considered the subjectivity of murderous and thieving maxims. The contingency of generating a contradiction can also be shown in terms of social and historical conditions. Suppose a leader who aspires to a political position and who is also in a position of strength and physical security (e.g., he controls or is the head of a military faction) considers killing his *only* rival. If the agent is permitted to consider his contingent circumstances, then the universalization procedure will adapt to these considerations. Similarly, if an agent lives in a time and society where poverty is scarce and the distribution of wealth is effectively egalitarian, then the universalization procedure will adapt to these conditions such that the universalized maxim will – here and now – undermine the necessary conditions for action.

that allows for the generation of a contradiction. First, while the observation that we have no duty to pursue a particular vocation or activity is correct, it is equally correct that the relevant duty we do have – namely, a duty of self-perfection – does not itself correspond to any principle of *action*. Thus, the relevant maxims are not *maxims of self-perfection* but, rather, maxims of *actions* that will realize and meet our duty of self-perfection. Relatedly, this is precisely the kind of maxim – that is to say, a maxim of an action which satisfies or violates duty – that is under consideration and has been considered problematic by Kantian interlocutors. The maxim does not, for example, introduce a degree of specificity comparable to unacceptable maxims such as “play tennis on *Sunday*” or “lie to my friend *Larry*.” Ironically, Allen Wood’s example of a maxim that obviously *cannot* be universalized (which I considered in chapter 1; see note 19) takes this same form: “I will never work, but always live by exploiting the labor of others.” In this case the maxim is one not just of sloth (i.e., not to work very much or not to work diligently) but *never* working. To put this somewhat awkwardly but also in a way that parallels the CEO maxim, the agent wills to be something very specific: a non-worker.<sup>9</sup> Interestingly, what makes this maxim non-universalizable is precisely the same thing that makes the CEO maxim non-universalizable: the specific social position willed by the agent (whether non-worker or CEO) cannot be filled by everyone without undermining the conditions (the work of others or others doing different work, respectively) that makes occupying the specific position possible. Second, I have already considered the difficulty of construing maxim thinly in chapter 2; very briefly, if we were to remove the specificity of vocation and replace it with a generic content (such as “develop my talents”

---

<sup>9</sup> We may also convert the CEO maxim to match the maxim proposed by Wood: the CEO proposes the maxim “I will never *labor*, but always live by supervising and managing the labor of others.”

or “earn money effectively”) then we would face the problem of letting in impermissible ancillary principles, such as principles of slavery or slave-holding (see chapter 2, 30-31).

## 2. *The Logical Contradiction Interpretation and conditional maxims*

As I have said previously, much of my explication of Hegel's objection is made in terms of the Practical Contradiction Interpretation. However, there are two interrelated objections based in the Logical Contradiction Interpretation that are worth consideration. To briefly review, the Logical Contradiction Interpretation seeks to establish that the proposed maxim is logically or physically impossible. More technically, willing the proposed maxim is impossible because the action of which the maxim is the principle is *inconceivable*. False-promising is again illustrative: according to the Logical Contradiction Interpretation, if the false-promising maxim were universalized, then trust with regards to such promises (i.e., repayment promises) would disappear and the convention of repayment promises would cease to exist. However, we are simultaneously attempting to conceive of the agent making such promises even though such promises do not exist. As Korsgaard writes, “If universalizing a maxim makes the action proposed inconceivable, then, we can get a logical contradiction” (1996, 82).

Korsgaard further contends that the advocate of the Logical Contradiction Interpretation can readily rebut Hegel's excessiveness objection:

It is true that we cannot imagine a world in which people give to the poor and there are no poor. Since there is no one to give to, it is an impossible state of affairs. But the advocate of the Logical Contradiction Interpretation can handle the objection. He can say that [one] has misstated the maxim. The maxim is to succor those who *need it*, and this maxim can be held in a world where no one needs help. The policy of succoring those who need it when no one does is not inconceivable. It merely gives one nothing to do. (Korsgaard, 1996, 87)

In one sense this is an odd passage since the maxims, on the face of it, are not substantially different; but Korsgaard reformulates the maxim to show how the advocate of the Logical Contradiction Interpretation can respond the “misstated” original formulation. Korsgaard’s preferred re-statement emphasizes those *who need it*; however, in this context, this is equivalent to the *poor* and if the latter generates a contradiction, so does the former. However, Korsgaard thinks the former formulation generates a contradiction while the latter formulation does not. Thus, Korsgaard must think there is a substantive difference between these maxims. To understand what this substantive difference is we have to answer the question: How might the maxim to “succor those who need it” be held in a world where no one needs help?

On the one hand, an answer is simple: the maxim is can be held conditionally. In other words, we help the poor when there is a need; and since, under the universalization of the maxim to help the poor there is no need, we simply do not act on the maxim though we still hold the maxim – we “have nothing to do.” It is important to see why this is possible with the maxim of generosity. Contrasting the maxim of beneficence with the false-promising maxim is helpful: The difference between these two cases lies in the fact that in the case of false-promising the conditions that are annihilated are not only conditions of action, but also conditions of conception: we cannot conceive of someone performing the action of a false-promise when such actions do not exist. In contrast, Korsgaard contends that the maxim to succor those who need it “can consistently be held (and in a degenerate sense acted on) in a world where no one needs it” (1996, 87).

I think this proposed solution is problematic. First, Korsgaard has not clarified how, in the absence of poverty, an agent with the maxim of generosity could

simultaneously have nothing to do *and* be able to act on this maxim even in a degenerate sense. I think he must mean that, in the case of universal false-promising, the agent could not take such an action because the existence of the action depends on a convention which is impossible given the state of affairs of universal false-promising; whereas, in the case of generosity, the agent could act – independently of any convention – if the need arose. This difference trades on the different grounds of impossibility and the different nature of the actions (conventional and natural, respectively). Second, the argument appears to elide the possibility of *acting* on a maxim and the possibility of consistently *holding* a maxim. Korsgaard appears to think that the false-promising maxim cannot be consistently held if the *conventional grounds* (*viz.*, trust in repayment promises) for acting on the maxim do not obtain. However, it is unclear why this would be a peculiar feature of conventional grounds not existing and not, rather, a feature of the necessary conditions for any action (such as helping the poor) not existing. Importantly, we cannot appeal to the inconsistency of *willing* the means to our end and simultaneously willing that these means be annihilated, as in the case of false-promising; to do so would be to move away from the Logical Contradiction Interpretation toward the Practical Contradiction Interpretation.

However, I want to put these initial problems to the side because I think there is a much more straightforward and serious problem; namely, the rebuttal is not generalizable to other examples. Korsgaard has rebutted Hegel's example of a maxim of generosity by contrasting it to a maxim of a conventional action; in doing so he has failed to appreciate that the rebuttal may depend on this difference and not on the success of the Logical Contradiction Interpretation. To see this, consider again the maxim of patriotism as well



as the maxim of self-perfection. In the former case we saw that the maxim is non-universalizable because if everyone were to join the military in defense on one's country, then one's country *as specified* would no longer exist, as well as one's enemies; and, further, one cannot defend what does not exist. As in the case of the false-promising maxim, the patriotic maxim cannot be consistently held because universalizing the maxim makes the action inconceivable: it is not merely that the patriotic has "nothing to do"; it is rather that the action for which the maxim is a principle has become impossible through universalization. The same impossibility holds for the maxim of self-perfection: if everyone were to will to become a CEO (and necessarily will the means to this end), then there would be nothing to be the CEO of; if there is nothing to be the CEO of, then it is inconceivable to take the action of being or becoming a CEO. Again, it is not merely that there is no current company for which the agent could be the CEO; it is that we cannot conceive of acting as a CEO within the context of CEOs institutionally not existing.

### 3. *The Practical Contradiction Interpretation*

We may now consider a more robust objection raised by Korsgaard in the context of defending the Practical Contradiction Interpretation. Korsgaard argues, in response to Hegel's objections, that according to the Practical Contradiction Interpretation the "the world of the universalized maxim only contradicts one's will if it thwarts one's purpose," (Korsgaard 1996, 95) but since one's purpose in adopting the maxim to "help the poor" is the supersession of poverty, the universalization of this maxim actually satisfies this purpose in a better way: poverty is entirely eradicated, which the agent alone could not

achieve. In this way, there is no *practical* contradiction, even though the means to one's end are annihilated, because the end itself is even more fully realized. This rebuttal would also appear to apply to our sympathetic agent's maxim to adopt a child: if all children were adopted, then the agent's end of alleviating the suffering and neglect of needy children would be even better realized than it would be through his own actions. Thus, if we accept the Practical Contradiction Interpretation, Hegel's excessiveness objection is ostensibly rendered moot.

The first difficulty with this rebuttal is that it does not respond to Hegel's objection in full; that is, to every implication Hegel identifies as following from the universalization of the maxim. Indeed, in responding to the excessiveness objection Korsgaard neither responds to the abstract structure of the objection, nor to Hegel directly. Rather, Korsgaard limits her response to the single example of helping the poor and to F. H. Bradley's (1927) reconstruction of Hegel's objection. Importantly, Bradley's reconstruction only includes the *absence* of poverty as the result of universalizing the maxim of beneficence. As we have seen, Hegel includes both the absence of poverty and universal poverty (or the absence of sufficient affluence) as possible consequences of universalization. Accordingly, it is not strictly true, when considering Hegel's actual objection, that the end adopted by the agent in his maxim is better realized by the universalization of his maxim to help the poor since the resultant state of affairs may be that *all* are poor. This is not a minor discrepancy: If generating a practical contradiction depends on the peculiar social conditions which obtain (i.e., if universalizing the maxim to "Help the poor" generates a contradiction when all would become poor, but not when poverty is annihilated), then the permissibility or impermissibility of an identical maxim

will be contingent on these conditions. That is, an identical maxim may be impermissible for one agent and permissible for another depending on the (expected) consequences of universalization. Consequently, an agent's duty (with regards to any maxim of this sort) would not be grounded in reason alone, but in particular and contingent empirical conditions.

In addition to this interpretive discrepancy, it is important to ask in what sense Korsgaard is right in her contention that the agent's actual end is realized by the universalization of his maxim to help the poor, even if the result of universalization is the absence of poverty, and whether this point generalizes to other maxims or is peculiar. We can answer this question by considering other canonical maxims. In the formulation of other maxims there is included an end which is situationally specific. For example, in Korsgaard's own example of an agent who considers a maxim of thievery, the agent assumes as his end *securing the property stolen*; he does not assume as his end *material security* in general. Likewise, in Kant's own formulation of the false-promising maxim, he attributes to the agent the assumed end of "getting out of a present difficulty" not the end of financial well-being generally (see Kant's *Groundwork* 4:402). Extrapolating from these examples, the proper ends belonging to maxims are, generally, situationally specific; and this contrasts with the end realized by universalizing the maxim to "Help the poor," which takes the eradication of poverty in general as its end.<sup>10</sup>

---

<sup>10</sup> This situational specificity rebuttal also addresses a parallel objection raised by Allen Wood (1990, 160) in that we may further distinguish possible ends adopted by the agent. Wood argues that the beneficent agent may adopt either the end of abolishing poverty as far as possible or the end of engaging in the activity of abolishing poverty. If the former is the agent's end, then there is no contradiction in the agent's will because poverty has been entirely abolished. However, if the latter is the agent's will, then according to Wood, the agent's maxim is already morally suspect. I am suggesting a third alternative: a maxim which has as its end the abolishment of poverty – or some significant part thereof – in the agent's particular situation. This, I believe, is the more plausible and,

This situational specificity may not make a difference when evaluating select maxims. For example, regarding the “help the poor” maxim, if all poor are helped and the poverty of all is relieved, then the poverty of the particular poor one intended to help is also relieved. Similarly, if we accept the first interpretation concerning the consequences of universalizing the patriotic maxim – that there are no agitators and thus no possibility of defense – then the agent's end of ensuring security for his country is even more fully realized. In these cases, and others, the satisfaction of one's situationally specific maxim corresponds to one possible consequence of universalization. However, this correspondence may be peculiar; a point punctuated by Hegel's claimed extension of his objection to an infinite number of maxims. If, for example, we accept the second interpretation of universalizing the patriotic maxim (that there will be either victory or defeat), then it does not seem to be the case that one's situationally specific patriotic end is even better realized. At best, it would seem that one's end of protecting and preserving one's country *may* be realized *if* one's country is the victor; but this again would make my duty a matter of historical contingency and not a product of pure practical reason. At worst, it would seem that by universalizing his maxim, the agent is putting his country at risk of total annihilation which is hardly consistent with his situationally specific end of preservation. The third interpretation is, again, most effective: if one's country (and enemies) as specified no longer exists, then one has clearly not more fully realized one's

---

at least, more common maxim concerning help for the poor. Few agents actually adopt the end of abolishing poverty *so far as possible*, even within their own community, but rather take up their duty of abolishing poverty where and as they encounter it. Furthermore, this situational specificity of ends accords with the formulation of maxims by both Korsgaard and Kant. As with Korsgaard's objection, this further, situationally specific option may not make a difference in the case of the “help the poor maxim,” but recognizing, in general, the situational specificity of ends adopted in maxims is relevant to any general solution to the excessiveness objection.

end of defending one's country. Thus, in order to assess the success of Korsgaard's solution we must first work through various maxims and the relevant contexts to see if the correspondence between situationally specific ends and the ends realized through universalization generalizes; an initial assessment of the maxims proposed here suggests that it does not.

Lastly, in evaluating Korsgaard's rebuttal we may also consider what consequences this augmentation of the universalization procedure might have for maxims generally by applying the same standard of success to maxims which would violate a negative duty. For example, instead of "help the poor" an agent considers the maxim to "kill the poor." To conform to the requirements of a situationally specific formulation, we may say the intention of the agent is to ease the tax burden on the rest of society and to increase overall tax revenue for other purposes, such as education and the environment, etc. In effect, this maxim represents a kind of *modest tax proposal*. Ordinarily – which is to say, prior to Korsgaard's rebuttal – we could reject this maxim by pointing out that the universalization of it would undermine the means for employing it: if all the poor were killed there would be no poor to kill. However, applying Korsgaard's standard of success in accordance with the Practical Contradiction Interpretation, this is no longer a reason for rejecting the maxim insofar as the universal eradication of the poor more effectively realizes the agent's end. Allowing a better realization of the agent's end to be included as a standard for passing the universalization test, then, would appear to have dire consequences for the validity of the test as a whole – or at least, the Practical Contradiction Interpretation.

## II. THE 'MURDERER AT THE DOOR' REVISITED

The excessiveness objection is an iteration of the false-positive objection, though with a specific scope and unique arguments. The 'murderer at the door' example also presents the challenge of false-positives: both objections argue that an action that intuitively ought to be permissible or is even obligatory is proscribed by the universalization procedure; the difference lies in the focus of the excessiveness objection on positive duties. I will first review Constant's objection and then turn to solutions proposed by Korsgaard and Wood.

The 'murderer at the door' example needs only a brief review. Constant accuses Kant of being committed to telling the truth "unconditionally" such that "it would be a crime to lie to a murderer who asked us whether a friend of ours whom he is pursuing has taken refuge in our house" (quoted in Kant, 1996, 611; Akademie edition 8:426). The counter-example mirrors the false-promising scenario which exemplifies the function of the universalization procedure: when we consider the proposed maxim of lying to the murderer at the door, the universalization of this maxim generates a clear contradiction: if everyone in the same situation and with the same end (the end is not made clear, but we may stipulate it is either "protecting a friend from extreme danger" or "frustrating the perpetration of a murder"), then the murderer would, just as in the false-promising case, scoff at our lie and conclude that we are indeed hiding our friend. And yet, the consensus is that not lying to the murderer would be terribly wrong, and perhaps we are even obligated to do so.

### A. Korsgaard's Response: Antecedent Evil

Korsgaard proposes a solution to Constant's objection that denies that the lie to the murderer at the door generates a contradiction. He develops this view by contrasting the murderer's attempted deception with the context and function of "lying in general." A false-promise functions because it is believed and the promisee trusts that the borrower will repay the money. In turn, the promisee trusts that this will happen and believes the liar because this is what people usually do. (Korsgaard is not appealing to what I have called – in chapters 1 and 2 – a normative expectation; but rather what might be called a *statistical* expectation.<sup>11</sup>) This expectation, according to Korsgaard, has implications for the permissibility of maxims: "But in most cases lying falls squarely into the category of the sort of actions Kant considers wrong: actions whose efficacy depends upon the fact that most people do not engage in them, and which therefore can only be performed by someone who makes an exception of himself" (1996, 136).

This expectation changes in the context of a murderer at the door, and this change in expectation has implications for universalization. Korsgaard posits that when a murderer comes to your door he must suppose that you do not know what he is about if he is to be successful. Simply, in this already very rare context, the murderer assumes that you do not know he is a murderer; and if he thought you did know he was a murderer he would conduct his business differently because people are generally not accommodating to murderers. This supposition about the knowledge of the person lied to has critical implications for the possibility of universalization: lying to the murderer at the door is permissible "because the murderer supposes you do not know what circumstances you are in – that is, that you do not know you are addressing a murderer – and so does not

---

<sup>11</sup> My thanks to George Sher for this terminology.

conclude from the fact that people in those circumstances always lie that you will lie” (Korsgaard, 1996, 136). If it is the case, upon universalization, that everyone will lie to a murderer and if a murderer knows that you know he is a murderer, then you cannot lie to him. However, the second antecedent is not true: the murderer does not know that you know he is a murderer, and thus does not conclude that you are acting according to the universal principle of lying to murderers. As such, your maxim is universalizable.

I think Korsgaard has identified something very interesting and important about the Formula of Universal Law and universalization procedure; namely, that the success of the universalization procedure depends on a particular epistemic criterion. I think that the epistemic criterion that must be met is one of *transparency*. Korsgaard has shown that if epistemic transparency is not a part of the murderer at the door example, then the deceitful maxim is universalizable. While Korsgaard is right in her analysis of the case and identification of an epistemic criterion, I do not think he appreciates how satisfying this criterion complicates the universalization procedure generally.

First, however, I want to consider the limitations of Korsgaard’s solution. That transparency is a necessary condition of generating a contradiction in the murderer at the door case is clarified by Korsgaard’s reluctant concession that a maxim of deceit could not be universalized in the case of an *honest* murderer. If a murderer comes to your door and says, “I’m here to murder your friend; is he here?” you cannot lie to this murderer without generating a contradiction. This is clearly problematic, and Korsgaard seeks to evade this difficulty by suggesting that “a murderer who expects to conduct his business by asking questions must suppose that you do not know who he is and what he has in mind” (Korsgaard, 1996, 136; emphasis is mine). In other words, Korsgaard supposes



that the practical efficacy of finding one's intended victim by asking questions depends on supposing the ignorance of your victim's accomplices. This supposition is clearly false: the SS officer or the *mafioso* who interrogates a terrified but courageous person giving refuge to the innocent may not only suppose that the person knows who they are and what they have in mind; they may actually rely on this knowledge and the fear and cowardly prudence it produces. And yet, the Formula of Universal Law would still proscribe lying in such circumstances.

Though limited, this solution may nonetheless be considered a small victory insofar as Kant is not as rigorous as his detractors have portrayed him, and the Formula of Universal Law is more adaptive than they have supposed. This conclusion is too hasty. In order to assess the success of Korsgaard's solution we need to consider what implications it has for the universalization procedure generally. Korsgaard has argued that the epistemic conditions which are ordinary for most cases of promising are not present in the case of the murderer. I think this is true; however, the same absence of epistemic conditions is true in a variety of deceitful scenarios. Take for example the con-artist. In order to conduct his business the con-artist must ensure that the "mark" is never aware of the confidence trick being perpetrated; and if the con-artist wills to be successful in securing his ill-got wealth, securing more wealth, as well as avoiding prison, he must also ensure that the mark *never* finds out. In short, the con-artist operates under *opaque* epistemic conditions. While confidence tricks are not ordinary in the sense of being perpetrated by most people, they are far too common and fundamentally contrary to Kantian principles: the con-man exemplifies the person who makes an exception of

himself.<sup>12</sup> And yet, as in the case of the murderer at the door, in the case of the con-artist the criterion of epistemic transparency is not met and the con-artist's deceitful maxim is universalizable: under the supposition that the mark does not know he is being conned the con-artist can conclude that his lie will not be ridiculed but will be believed – at least as much as in any ordinary circumstance. Korsgaard, then, has correctly identified a crucial epistemic criterion at work in the murderer at the door example but has not appreciated the variability of this criterion or the implications ensuring epistemic transparency has for the success of the Formula of Universal Law generally.

In short, in the case of the murderer at the door the perpetrator of evil supposes that you do not know of his nefarious scheme and this supposition is *incorrect*; in the case of the con-artist the perpetrator of evil supposes that you are not aware of his nefarious scheme and this supposition is *correct*. Thus, in order for Korsgaard's solution to succeed some justification must be given for why we must attribute epistemic transparency – that is, knowledge of nefarious intentions – to the victim in all cases of nefarious action. Without this attribution of transparency, the original objection to the universalization procedure actually becomes more severe in that it now includes the

---

<sup>12</sup> I have chosen the case of a con-artist because it represents a relatively common practice while also exemplifying the difficulty of removing epistemic transparency. However, once we identify epistemic transparency as a necessary condition for generating a contradiction *and* allow for the manipulation of this transparency (i.e., once we allow the respective agents to know or not know the motivations or assumptions of the other), problematic cases multiply, including Kant's standard cases. For example, if we remove epistemic transparency from the original false-promising example, it is not clear that we can still generate a contradiction. This example depends on the victim knowing the motivations and financial need of the liar as well as the same epistemic transparency holding for all other cases of false-promises by those in financial need. If the fact of financial need can be kept from the victim – if the circumstance remains epistemically opaque – then promising in general would continue and one would never know if they are being given a false-promise. Perhaps false-promising would become common enough that people would be consistently suspicious of repayment-promises, but perhaps not; either way, the non-universalizability and consequent impermissibility would then rely too heavily on the particular social conditions of the agents, and not the exceptional character of the maxim.

generation of false-*negatives* in addition to false-positives.

## B. Wood's Exegetical Rebuttal

In his recent work dedicated to Kantian ethics, Allen Wood (2008) takes a very different approach: Wood does not propose a solution to Kant's rigorism, but rather argues that "Kant's views about veracity are reasonable or at least defensible, if not self-evident" (240). Wood's defense takes two forms: on the one hand he argues that Kant's apparent strictness can be attributed to a "warranted rhetorical exaggeration"; on the other hand, he argues that the infamous response to Constant's murderer at the door example has a limited scope and that Kant's position has been maligned because it is misunderstood. In this way, Wood proposes an exegetical defense of Kant.

Turning first to the murderer at the door objection, Wood contends that there is a lack of understanding concerning the *kind* of falsehood under consideration in Kant's response to Constant. Wood clarifies the dialectical context by giving a taxonomy of falsehoods, and arguing that Kant's reply only concerns one species in this taxonomy: namely, false *declarations*. First, on Wood's interpretation, a *lie* is analytically impermissible: a lie is "an intentionally untruthful statement *that is contrary to duty*" (Wood, 2008, 240). Thus, Kant's unconditional proscription of lies is trivially true insofar as such a proscription is tautological. What is of concern, then, are not lies but what Wood translates as "falsifications" (*falsiloquium*) and which falsifications are lies.

According to Wood, Kant's considers falsifications as lies *only* in terms of two different kinds of duty: duties of *right* and perfect duties to oneself (Wood, 2008, 241). We must note immediately that this is a peculiar and confusing claim. Some insight into

what Wood may mean here can be gained by considering a succeeding footnote: Wood observes (324, note 3) that in a variety of works, including the *Metaphysics of Morals*, his lectures on ethics, and the essay on the right to lie, Kant “never discusses lying as a violation of a duty to others by relating it to the formulation of universal law or the formula of humanity.” It may be that Wood considers all false-promises to be impermissible declarations, as Kant categorizes them in *The Doctrine of Virtue* (*MM*, 6:429), and has concluded that false-promises should only be understood as violations of a duty of right. But this would, inexplicably, omit from our taxonomy of duties the two sections in the *Groundwork* (4:422 and 429-430) where Kant does discuss lying, in the form of false-promising, in terms of a violation of our perfect duties to others. Moreover, Wood has explicitly recognized in his earlier work that false-promising, which Kant explicitly identifies as a kind of false declaration, is a form of coercion and deception, and that these forms of action “obviously violate Formula of Humanity because they achieve their end precisely by frustrating or circumventing another person’s rational agency and thereby treat the rational nature of the person with obvious disrespect” (1999, 153); in other words, they violate a perfect duty to others. Thus, to claim that Kant, in his canon as a whole, considers lying – that is, impermissible falsifications – only in terms of duties of right and perfect duties to oneself is textually unsupportable.

As for Kant’s circumscription in the *Metaphysics of Morals* and the essay on the supposed right to lie: first, as we shall see shortly, Wood’s own explication of what is at issue in the latter essay explains the limited consideration given to lying; regarding the former text, if we were to take Kant’s taxonomy and explication of duties in the *Metaphysics of Morals* as representing all the general duties we have, then we would have

to conclude that we have *no* perfect duties to others that are also ethical duties except ethical duties of respect such as that against *mockery*, insofar as Kant considers only perfect duties to oneself *qua* animal being and *qua* moral being, and then only *imperfect* duties to others. This would be directly contrary to Kant's emphatic contention that "all duties, just because they are duties belong to ethics" and that what distinguishes a duty of right from an ethical duty is the kind of incentive, internal or external, that may be attached to it (*MM*, 6:220).

Turning to Wood's explication of Kant's response to Constant, Wood argues that in the essay on the supposed right to lie Kant is considering falsifications only *qua* violations of duties of right. Crucially, violations of right are not understood in terms of the Formula of Universal Law (or the Formula of Humanity or the Formula of the Kingdom of Ends) but rather a separate principle: "Any action is right if it can coexist with everyone's freedom in accordance with a universal law, or if on its maxim the freedom of choice of each can coexist with everyone's freedom in accordance with a universal law" (*MM*, 6:230). Importantly, it is not the case that *any* falsification violates this principle; neither is it the case that a falsification that violates the Formula of Universal Law (or the Formula of Humanity) is a violation of this principle and thus contrary to a duty of *right*. Thus, in order to determine when a falsification becomes a lie in terms of right, we have to determine under what conditions this principle of right would be violated by a falsification.

At least one set of falsifications that violate the principle of right are false *declarations*. Declarations are another technical term in Kant and "refer to statements that occur in a context where others are warranted or authorized in relying on the truthfulness

of what is said, and make the speaker liable by right” (Wood, 2008, 241). According to Wood, oaths (both in juridical contexts but also religious contexts) as well as statements made within a contractual agreement exemplify declarations. A false declaration violates the principle of right because it infringes on the freedom of others; that is, universal false declarations cannot coexist with everyone’s freedom. Importantly, it is not merely the person to whom I make a false declaration that is affected by my lie; once a practice is instituted that establishes right, such as a practice of contract or legal testimony, my false declaration affects everyone who’s free action depends on this practice.<sup>13</sup> It is the particular and essential social function of declarations in ensuring freedom of choice that grounds Kant’s unconditional prohibition on false declarations.

With these exegetical points established Wood argues for the reasonableness of Kant’s position by highlighting the artificiality of Constant’s example and the necessary conditions that would obligate a person to be honest with a murderer. First, a murderer at the door must have the right to demand or expect a declaration from you; in Constant’s example the murderer has no such right. It may be the case that if a police officer approached your door, or another authorized agent of the state, then you could be obligated to give a declaration. However, without the required context of a legal investigation, proceeding, or contractual agreement Kant maintains that one is under no obligation to give a declaration, and thus you *may* permissibly give a falsification without violating a duty of right. Second, Wood argues that Kant defends the concept of the

---

<sup>13</sup> “If someone lies in a court of law, for example, it is not only his adversary whose right is violated but the entire system of right, which must presume the truthfulness of declarations made in legal processes. If someone is defrauded in a contract, it is not only this person whose right is violated but the entire system of contract right.... By making such a declaration, I am in that sense acting in a way that, if its permissibility were generally allowed, would deprive all declarations of their validity, whether or not I intend that result and whether or not it actually occurs” (Wood, 2008, 243).

‘necessary lie’: “Kant explicitly allows that no lie, and no violation of right, occurs if we commit a falsification in order to prevent another from making a wrongful use of the truth” (Wood, 2008, 246). Wood’s justification of this exception is drawn entirely from Kant’s *Lectures on ethics*:

Again, I may make a false statement (*falsiloquium*) when my purpose is to hide from another what is in my mind and when the latter can assume that such is my purpose, his own purpose being to make a wrong use of the truth....If force is used to extort a confession from me, if my confession is improperly used against me, and if I cannot save myself by maintaining silence, the my lie is a weapon of defence. (*LE*, 228)<sup>14</sup>

In sum, then, a falsification is only *absolutely* prohibited if three conditions are met: first, the false statement is a declaration; second, the other person must be in a position that gives him warrant to believe the statement; and third, the false declaration is not an instance of a necessary lie. With these conditions, Wood concludes that Kant’s position is (for the most part) reasonable, and that the attribution of excessiveness strictness depends on an “extreme and artificial” framing of the question.

Wood’s explication is a welcome and needed clarification of all the issues involved, and certainly complicates the dialectic; nonetheless, it is not entirely exculpatory. This may be seen in a variety of ways. First, even while limiting the discussion to declarations and violations of right and conceding the three conditions Wood has explicated, Kant’s view remains unacceptably stringent, even if less so than previously thought. Wood has given oaths and contracts as representative examples of

---

<sup>14</sup> Wood quotes from the Heath translation: “I can also commit a *falsiloquium* when my intent is to hide my intentions from the other...since his own purpose is to make a wrongful use of the truth. ...So far as I am constrained, by force used against me, to make an admission, and wrongful use is made of my statement, and I am unable to save myself by silence, the lie is a weapon of self-defense. (*LE*, 27:447-448)

declarations; however, as noted above, Kant also includes *promises* in this category, and this designation accords with the distinguishing criteria: promises are a kind of statement that function within a social practice, provide warrant for believing a statement, and also make the giver of the promise liable. Moreover, with the establishment of this practice, false-promising violates the principle of right that determines the impermissibility of an act in terms of duties of right; specifically, as in the case of a false contract, one cannot make a false-promise without wrongfully restricting (Wood, 2008, 243) the freedom of others.<sup>15</sup> Thus, though the case Constant presented may be excessively artificial, other cases which satisfy the three conditions yet remain problematic are not. Indeed, we need only provide a case where a promise is given, or must be given to provide refuge to an innocent person, in order to satisfy the first two criteria and revitalize the problematic case.

This leaves the third criterion: that the declaration is not an instance of a necessary lie. Exegetically, I think the inclusion of this criterion dubious in that it relies exclusively on Kant's lectures and finds no corroboration in Kant's published works. More importantly, Mahon argues (2006, 672ff) that even if the criteria for a necessary lie are met, the prohibition on lying remains excessively restrictive, and that a maxim formulated in terms of some of the conditions which permit exceptions to the duty against lying (e.g., that one may lie when under threat of violence or to thwart the immoral use of

---

<sup>15</sup> An example from current law may help to support this point. In a recent decision, the California Supreme Court held that companies can be held liable for false assurances made to prospective employees: "If a worker can prove that he or she accepted a job under false assurances of security, employers can be held liable for potentially whopping punitive and emotional distress damages that aren't normally available in termination cases." *Justices Say Employers Can Be Liable for False Promises* [Online.] Available at: [http://articles.latimes.com/1996-01-30/news/mn-30329\\_1\\_employment-lawyer](http://articles.latimes.com/1996-01-30/news/mn-30329_1_employment-lawyer) (accessed March 22, 2017).



the truth) are non-universalizable. However, as we do not need to satisfy this criterion to produce a contradiction, I will put this interpretive point to the side. To see how Kant's position remains excessively stringent, we may consider two cases: the first is Wood's own example of an unscrupulous prosecutor; the second is an example of a corrupt, murderous police office.

In an effort to show the reasonableness of Kant's position once the context and conditions have been made clear, Wood asks us to consider the case of providing testimony under oath when an unscrupulous prosecutor is prosecuting a man – your friend – whom you know to be innocent. Strangely, Wood seems to abandon one condition which permits a necessary lie, that the truth will be used for immoral ends, when considering this case – an unscrupulous prosecutor would seem to ground an exception. Moreover, Wood does not actually *argue* that in such a case you would be obligated to give an honest testimony, but merely states that *his position* is the Kantian one; and importantly, this is because telling a falsehood even in such a compromised context would make you “the one turning the process into a sham, by behaving according to a principle that, if generally followed, would bring all solemn testimony and all legitimate legal processes into discredit” (Wood, 2008, 249).

This is not a compelling response. It must be emphasized that Wood has not provided any reason other than the same contested Kantian reason for why there should not be an exception in this case. However, the consequences of universalization – even in the murderer at the door case – are not being contested; what is contested is the claim that the consequences of universalization – good or bad – should serve as a guide for what we ought to do. To appeal to such consequences of universalization to justify the

unexceptional Kantian position does not move the dialectic forward, it returns us to the beginning. This is perhaps uncharitable, as it may not represent the strategy in this passage. Rather, the strategy, perhaps, is better understood as presenting a more realistic case with all of the conditions met that has strong intuitive appeal and, thereby, shows (but does not *demonstrate*) the reasonableness of Kant's position. However, this approach is also unsuccessful. First, this reader, at least, finds the Kantian position counter-intuitive: if you *knew* that your friend was innocent, and you also *knew* that the prosecutor was unscrupulous (I take this to mean that the prosecutor was knowingly prosecuting an innocent man with the intention of killing him or some other immoral end), and the legal system was generally legitimate, it would still (in my mere opinion) be right to lie: I would feel overwhelmingly obligated to lie for my friend, and I would expect my friend to lie for me in the same situation. However, conflicting intuitions get us nowhere, and are utterly contrary to the Kantian project.

What may be useful is analyzing what accounts for the claimed intuitiveness. In key respects, Wood has not provided a scenario which parallels the murderer at the door case. Lying under oath is a serious act, and a serious violation of norms established for the preservation of justice. As such, many will find any such violation intuitively impermissible if it is not necessary. A key feature of the murderer at the door scenario is the assumed fact that the person proposing the lie has no other option – that is, he faces a genuine dilemma. In contrast, in the circumstance of considering false-testimony there is much that one could do – one could vocally protest (inside or outside the court) and call attention to the unscrupulous prosecutor, one could plead the 5<sup>th</sup> (even if testimony would not self-incriminate), one could go to the media, one could protest in other ways (e.g., a

hunger strike) and attempt to stall the proceedings. Similarly, this case does not possess the same immediacy of the murderer at the door case. In the latter, the assumption is that a lie *at this moment* and *by you* is necessary to avoid the death of the potential victim. In contrast, in the case of a legal proceeding there remain additional steps that one may take to protect the innocent, and given the complex nature of a prosecution, defense, and judgment the responsibility for protecting the innocent is held corporately. These disanalogies may skew our intuitive response. Thus, Wood has provided an additional case, but has not provided the exigency which characterizes the murderer at the door case, and which undergirds the intuition that it is permissible to lie and one may even be obligated to do so.

A better example is one also mentioned, albeit briefly, by Wood. Suppose a police officer comes to your door asking after your friend. Just as in the case of the stranger who seeks to murder an innocent person, this officer is corrupt, unscrupulous and you *know* that he is seeking to kill your friend who is innocent. In order to satisfy the condition that the false statement constitute a declaration, we may stipulate that in an effort to placate the suspicious officer, you promise that you are ignorant as to the person's whereabouts. This example has several important features: First, though such a scenario is admittedly rare, contrary to Wood's exegetical rebuttal it is not artificial; this is a scenario that both satisfies all of the conditions necessary for prohibiting a false declaration while also being plausible. Second, unlike the example of false-testimony, the need for a lie in this context is urgent, both in terms of immediacy and the lack of other options (one could refuse to answer the police officer, but this would likely raise suspicion). Third, as in the case of the unscrupulous prosecutor, the example does not satisfy Wood's requirement for

a necessary lie: though the officer is using your testimony for an evil purpose (as is the prosecutor), you are not being forced – through violence or threat of violence – to tell the truth; you could keep silent or simply close the door, but judge that this is not the best or right action for the protection of your friend. Lastly, I suggest that the example has an intuitive force comparable to the murderer at the door example (i.e., in this scenario it would be permissible and possibly obligatory to give a false declaration); and yet, the case directly and clearly violates the principle of right, which may be seen by simply interpolating “promises” into Kant’s own explication of this principle: By making a false-promise “I bring it about, as far as I can, that declarations in general are not believed, and so too that all rights which are based on [promises] come to nothing and lose their force” (8:426).

The foregoing response to Wood has been in terms of duties of right; this limitation conforms to Wood’s contention that Kant’s response and rigorism is relevant to declarations. We need not, however, understand Constant’s objection in such narrow terms, and once we remove this interpretive limitation we may see that the response is wholly unconvincing when considered in the context of the violation of ethical duties to others and the Formula of Universal Law. This is, for our purposes, the most important and simplest rebuttal to Kant’s conditional prohibition of false declarations: namely, that there is no reason to maintain this historical and textual circumscription. We may concede Wood’s exegetical point – that Kant’s historical reply to Constant was limited to declarations – while recognizing that the example of the murderer at the door has been appropriated into a discussion of Kantian morality – that is, ethical violations of perfect duties to others. The fact that this appropriation was initially based on a misinterpretation

is moot. In other words, we can still ask, “How would Kant and how can Kantians respond to the case of murderer at the door in the context of perfect ethical duties?”

Wood recognizes that Kant’s own position regarding the perfect duty to be truthful “is far more extreme than anything Kant says in the essay on the right to lie. Taken literally, he is saying that any intentional untruth whatever violates a duty to oneself” (2008, 252). There are two crucial points to be made regarding Wood’s response to this apparent ethical rigorism: first, as noted above, Wood inexplicably restricts our perfect duty not to lie to perfect duties to oneself; but a consideration of all of Kant’s works shows that this circumscription cannot be maintained: false-statements, as exemplified by but not limited to false-promises, are clearly a legitimate target of the Formula of Universal Law (and Formula of Humanity) as potential violations of perfect duties to others. Second, Wood does *not* think we should take Kant’s absolute and stringent prohibition on untruthfulness literally and furnishes textual evidence for Kant’s hyperbolism. This proposed hyperbolism, I contend, faces both exegetical and theoretical challenges.

The textual evidence furnished by Wood is limited and problematic. Wood provides three sources in support of Kant’s “rhetorical exaggerations”: The first source comes from Kant’s casuistical questions in *The Doctrine of Virtue*. The second form of support comes from Kant’s lectures on the metaphysics of morals, where he observes that a “moral casuistic would be very useful...if the limits were defined as to how far we may be authorized to conceal the truth without detriment to morality” (from *Vigilantius’s lecture notes* 27:701; quoted in Wood, 2008, 253). The third source, also from Kant’s lectures, consists in another appeal to the concept of necessary lies. Exegetically, an

appeal to Kant's *Lectures on ethics* provides weak support. These lectures, compiled by students, were given between 1775 and 1780 five years before the publication of the *Groundwork* (1785) and a full seventeen years before the *Metaphysics of Morals* (1797). The prior aspiration for a casuistic that would provide the boundaries of latitude for moral action does not imply the consistency of such latitude with Kant's mature and considered position. This skepticism concerning the compatibility of intuitive latitude extends to Kant's putative support of necessary lies. Admittedly, this exegetical weakness does not prevent us from integrating such a casuistic that would establish justified latitude, so we must consider such integration and the possibility of latitude.

To this end, the casuistical questions in *The Doctrine of Virtue* are both more exegetically provocative and theoretically intriguing. Kant asks whether an untruth may be told for the sake of propriety: the first example is of the polite valediction, 'your obedient servant'; the second is of a lie given in order to avoid embarrassing another with your honest view of his work. Kant takes the impermissibility of these lies to be questionable because, in the former instance, no one is *deceived*, and in the latter instance, the polite lie is "expected of one" (*MM*, 6:431). It must be noted, as Wood himself does, that these exceptions are not part of Kant's formal position, but rather difficult cases presented for reflection. Kant has not, that is, endorsed either exception. Despite this, Wood claims that he has provided a theoretical means for accepting or justifying either exception. To reiterate the point made about latitude found in the *Lectures*, Kant's reluctance or inability to formally endorse such latitude here need not prevent contemporary Kantians from doing so; as such, they require our further attention.

It must be emphasized that the desire for latitude is not contentious (at least not

presently); what is contentious is the theoretical compatibility of such latitude. That is, even if we were to grant that Kant is amenable to latitude regarding a duty of truthfulness and that contemporary Kantians should be, it remains entirely unclear how this latitude may be systematically integrated into Kant's moral theory or the theory of his contemporary defenders. In other words, a casuistic *would* be very helpful, but is precisely the absence of a systematically tenable casuistic that continues to problematize the Kantian position. This is particularly true for the Formula of Universal Law.

We can see this difficulty in the examples given. The first casuistical question provides little help in absolving Kant's rigorism considered here in that the reason given for the latitude is that no one is *deceived*.<sup>16</sup> In contrast, in the murderer at the door case the murderer is deceived, and such deception is the liar's objective and what is considered permissible or even obligatory. Thus, making an exception for lies that do not deceive is ineffective: what is needed is an exception and justification for deliberate deception. In the second case, the latitude is justified by social expectation. This is a more difficult justification: if one may lie when it is expected of one, the converse may also be true, and one might argue that a lie is excepted when there is no legitimate expectation of truth. Much like Wood's analysis of Kant's position on declarations, this absence of a legitimate expectation would exonerate the agent who lies to the murderer at the door.

This solution has intuitive appeal: we should make exceptions in accordance with the presence or absence of legitimate expectations. The difficulty with this solution is the justification of what counts as "legitimate." While we may all agree that a murderer has no legitimate expectation to the truth, we may very well disagree with other examples,

---

<sup>16</sup> I discuss this apparent change of the criterion for impermissible falsehoods in more detail in the next chapter (see page 187ff).

including Kant's own example of lying about the quality of another's work, Wood's example of testimony under oath in the false prosecution of an innocent man, and innumerable other examples. Thus, Wood contends that we along with Kant will recognize that sometimes telling the truth is absurd – and no doubt this is true – but he provides no grounds for identifying such absurd truths except our own intuitions and social conventions; nor does he provide any way of consistently integrating these into Kantian morality generally.

Most importantly, thus far Wood has considered only the possible latitude Kant recognizes for perfect duties to *oneself* that are derived from the Formula of Humanity; he has not considered whether the same latitude may be extended to perfect duties to *others* derived from either the Formula of Universal Law or the Formula of Humanity in the *Groundwork*. However, what is of present concern – that is, what Constant's objection as appropriated within contemporary Kantian literature challenges – is the relevance of latitude in our derivation of duties from the Formula of Universal Law, and on this front there are no resources in Kant to suggest a procedural latitude, and there is no proposal for how either to integrate such latitude into the universalization procedure itself or to identify such latitude through the introduction of a superseding test or principle.<sup>17</sup> This point is of particular importance for the latitude putatively established with the concept of necessary lies. Admittedly, the conditions which make a lie necessary have intuitive appeal, and this appeal is particularly strong in the context of the Formula of Humanity –

---

<sup>17</sup> We must be charitable and recognize that Wood (1999 and 2008) has argued at length that the Formula of Universal Law is not Kant's final or fullest formulation of a practical law of reason: "Thus since FUL and FLN already presuppose these later formulations, neither of them can be regarded as an adequate and self-sufficient formulation of the supreme principle of morality" (1999, 81). However, Wood does hold that the Formula of Universal Law functions as a "permissibility test for maxims" and it is in this function only that we are defending Constant's objection.



it is difficult to accept that one is treating the murderer as a mere means by frustrating his nefarious scheme that would treat another as something that is literally disposable.

However, this intuitiveness poses a problem, rather than a solution, in the context of the Formula of Universal Law: How are our intuitions to function in relation to the universalization procedure? To follow Kant, though we have strong intuitions regarding necessary lies we still do not possess a rational casuistic which would enable us to integrate this intuition into our theory, and particularly that procedural element of our theory which tests for the permissibility of maxims. Without such a rational casuistic our latitude would be based entirely on intuition and, consequently, we would be theoretically committed to the superseding role of intuition (“What other intuitions might we now integrate?”) which would entail a fundamental departure from the Kantian project.

By way of summary, then, Wood has identified the uncontroversial need for latitude, and, more controversially, Kant’s aspiration for and amenability to latitude. However, in so doing he has also identified the limit of *Moralität* and the crux of Hegel’s objection: the moral standpoint alone cannot adequately furnish concrete duties because it cannot systematically provide or integrate content that justifies the required latitude – concretely, it cannot tell us what counts as a legitimate exception or what is expected of us. Neither can Morality *per se* provide a justification for deceptive lies even in the case of necessary lies (despite the latter’s intuitive appeal). In this way, the proposed defense of Kantian morality is a signpost, pointing the way to Hegel’s sublation (*aufgehoben*) of *Moralität*. In specific terms, Morality must become a theory of actual duty, fulfilled in an ethical life (*Sittlichkeit*), where the limits of required honesty, and the absurdity of punctilious honesty, among other duties, are understood by properly educated and facile

practitioners.

### III. CONCLUSION

The problem of *showing to much* is not new, as Contant's contemporaneous objection makes clear. The purpose of this chapter, then, has not been to raise a new objection, but rather to show the extent and continued relevance of this problem for the Kantian moral project. This has been accomplished by, first, providing an interpretation of Hegel's excessiveness objection that shows it to be a corollary of his emptiness objection and to have an abstract form that can be applied to a variety of maxims. This interpretation has shown how Hegel's objection is unique, and specifically how the problem of showing too much – or false-positives – is not merely a technical difficulty which may be resolved by resolving the question of how to formulate maxims; but is rather a problem for Formula of Universal Law and universalization procedure as such; in this way, the problem of showing too much has an extended scope. Furthermore, by moving away from interpretations of Hegel that focus on individual examples (such as the maxim "Help the poor"), the present exposition of Hegel's objection has been effective in responding not only to general rebuttals, but also apparently successful rebuttals grounded in both the Logical Contradiction Interpretation and Practical Contradiction Interpretation.

The goals of the chapter were also achieved through a reconsideration of the problem of the murderer at the door and two very different proposed resolutions to this canonical objection. Korsgaard has presented a very interesting solution to the problem, whereby antecedent evil accompanied by epistemic opacity on the part of the perpetrator (i.e., the murderer) enables the lying maxim to be universalized without generating a

contradiction. While conceding the technical analysis of the scenario, I have challenged this solution by considering both the limitations of the solution (which Korsgaard concedes) as well as the implications of manipulating this new condition of epistemic transparency when employing the universalization procedure. Identifying epistemic transparency as a condition of the success of the universalization procedure is a unique and important insight. However, allowing for this condition to be eliminated from our consideration of cases (i.e., introducing epistemic opacity into the scenarios) has dire consequences for the universalization procedure generally: if epistemic opacity were generally allowed, then the Kantian would face myriad cases of false-*negatives* in addition to the problem of false-positives with which we started and the analysis of epistemic transparency was supposed to resolve.

In a very different, exegetical approach, Wood argued that the apparent rigorism of Kant's response to the 'murderer at the door' example is attributable to the artificial features of the case, and if these features are removed and we limit Kant's response to the appropriate context of *declarations*, then his absolutist position appears (mostly) reasonable. In response, I argued, first, that even with all of the legitimate conditions for making declarations met, the Kantian position still faces strongly counter-intuitive cases – that is, the Formula of Universal Law still shows too much. Second, though Wood's contextualizing of Constant's objection to duties of right is of historical importance, this does not prevent the appropriation of his objection to the realm of ethical duties, and specifically ethical duties to others. In light of this appropriation I considered a second exegetical argument by Wood – namely, that Kant's absolute proscription of untruthfulness is a product of pedagogic hyperbolism. In response, I argued that even if

we accept the interpretive point, the Kantian position is still theoretically rigorous in that no proposal has been furnished for how to integrate latitude, regardless of its desirability, into the universalization procedure or Formula of Universal Law. In this way, rather than showing the actual reasonableness of the Kantian position, Kant's supposed hyperbolism and the Kantian desideratum of latitude reveals the limits of *Moralität* and the need for fulfilling the moral point of view with content-full ethical life (*Sittlichkeit*).

All of these arguments as well as those in chapter 1 and 2 have been limited, for the most part, to the Formula of Universal Law. Kantian interlocutors have long recognized the limitation and problems facing the Formula of Universal Law, and turn to the Formula of Humanity as a more substantive and successful iteration of the moral law. In particular, Kantians argue that Hegel's emptiness charge is limited to the Formula of Universal Law and, thus, is not an objection to the Kantian moral position as a whole. Accordingly, in the following chapters (4 and 5) I adapt Hegel's arguments, applying them to the Formula of Humanity and showing the broad relevance of the emptiness critique.

## Chapter 4:

### Hegel's Emptiness Critique, the Formula of Humanity, and Negative Duties

There are two forms of objection raised against Hegel's emptiness objection. The first is that Hegel has misunderstood Kant's arguments regarding the Formula of Universal Law, how this first formulation functions and is to be applied. I have addressed this objection directly in chapter 1, arguing for an alternative interpretation of Hegel's objection, and expanded on the implications of this alternative interpretation in chapters 2 and 3. The second objection raised against Hegel's emptiness objection is that Hegel, despite presenting his objection as applying to the *moral point of view* generally, focuses exclusively on the first formulation of the Categorical Imperative – the Formula of Universal Law – without directly applying his arguments to the further two formulations, the Formula of Humanity and the Formula of the Kingdom of Ends. This objection is of particular import insofar as the Formula of Humanity is ostensibly content-full: positing rational beings as *ends in themselves* and respect for their *humanity* as the end of all moral action. The putative fullness of the Formula of Humanity is also used as a foundation for solving the formalistic deficiencies of the Formula of Universal Law.<sup>1</sup>

---

<sup>1</sup> See, for example, Herman (1993e, 84ff): “Thus it does not seem possible for the RMS [rules of moral salience] to be generated by the CI [universalization] procedure. But if they are independent of the CI, the unity of the Kantian system is compromised, consisting partly of formal procedure and partly of something like preprocedural intuition or convention....Independence from the CI procedure, however, does not entail independence from the Moral Law....We want to be in a position to say that although the RMS are not and cannot be derived from the CI procedure, they are neither arbitrary nor conventional, for they express the same fundamental concept (the Moral Law) that the CI procedure

My critical aim, then, is equally clear: I must argue that the Formula of Humanity does not escape Hegel's critical gaze and that his emptiness objection, or rather critique, may be extended to these apparently content-full forms of the Categorical Imperative. Following the secondary literature, I will focus primarily on the derivation of duties from the Formula of Humanity, and then argue how the arguments leveled against the Formula of Humanity may, in some cases, be adopted and applied to the Formula of the Kingdom of Ends. Even so circumscribed, this objective requires an expansive discussion of both Kant's own texts as well as an extensive and varied Kantian secondary literature. First, Kant himself has different desideratum for the derivation of negative and positive duties; for the former he intends to derive particular, concrete duties, while for the latter he limits his derivation to general duties – such as self-perfection and practical love. The extension of Hegel's critique must be adapted accordingly, and I will develop Hegel's arguments for the derivation of negative duties in the present chapter, and for positive duties in the following chapter (chapter 5).

Even when considering the derivation of negative duties only, Kant provides very different arguments for the derivation of duties from the Formula of Humanity in the *Groundwork of the metaphysics of morals* and *The Doctrine of Virtue* in the *Metaphysics of Morals*. Moreover, Kantian proponents interpret Kant differently and offer competing derivations of particular duties, competing duties to be derived, as well as competing

---

represents for moral judgment." Herman contends, moreover, that conceiving of oneself as an equal moral agent provides the foundation for rules of moral salience as well as the foundation for judging other rules of moral salience, of other cultures, as inferior. Failure to "acknowledge" the moral status of other rational agents is sufficient grounds for moral criticism. The relevance of Hegel's arguments extended to the Formula of Humanity is clear. The extension provided in this chapter aims to show that acknowledgement of such an abstract status is not sufficient for distinguishing the permissible from the impermissible and determining an immanent doctrine of duty.

solutions to problematic cases. Addressing every interpretation of Kant's derivation of duties is not feasible here. As such, I limit my extension of Hegel's critique to the derivation canonically prominent duties, and limit my discussion of Kantian proponents to those that either offer a direct defense of the fecundity of the Formula of Humanity or offer an influential defense. The specific critical goal of the chapter is to defend and elaborate on Hegel's contention that no "immanent doctrine of duties" can be derived from the Moral standpoint. This is accomplished by either demonstrating that the derivation of any particular duty requires moral material from outside the Formula of Humanity or that the criteria set forth for judging when an action violates the Formula of Humanity are so broad or so strict that they are insufficient for distinguishing genuine duties from false duties. Put simply, the Formula of Humanity *alone* is insufficient for deriving particular duties or the Formula of Humanity is insufficient for getting the derivation of our duties *right*.

With these caveats and objectives in mind, the chapter proceeds as follows: In Section I I briefly provide textual evidence supporting the contention that Hegel at least intended his emptiness critique to apply to the Categorical Imperative generally, even if one must concede that he did focus too narrowly on the Formula of Universal Law in the application of his arguments. In the following section (Section II) I turn to arguments by Hegel's Kantian interlocutors that contend the Formula of Humanity is sufficiently substantial to provide determinate, concrete moral guidance. In other words, the Formula of Humanity is immune from Hegel's emptiness critique. This latter contention relies on an interpretation of Kant's Formula of Humanity, forwarded by his defenders, as (quasi) *teleological* and not strictly *deontological*.

In Section III I turn to specific arguments and interpretations concerning the derivation of particular duties from the Formula of Humanity. I begin with the derivation of particular duties from Formula of Humanity as explicated by Kant in *The Doctrine of Virtue*, and give particular prominence to Lo (1989) and Wood (1999) who specifically defend – with varying degrees of confidence – a syllogistic interpretation of these derivations. In response, I argue that in the case of *negative* duties (such as the duty not to engage in suicide or lying) such syllogistic derivations of duties require an ancillary syllogism in support of the intermediate premise. Moreover, such dependence on independent moral claims is clear from Kant's own explication. Thus, contrary to Lo's and Wood's explications, such ancillary arguments rely on content that is both logically and morally independent of the Formula of Humanity as well as highly contentious.

In the second half of Section III.A, I consider the strongest Kantian derivation – presented by both Korsgaard and Wood: the proscription against suicide and deception derived in the *Groundwork*. In response, I first reiterate my previous argument that sufficient criteria for distinguishing between using others as a means and using others as a *mere* means (and thus violating the Formula of Humanity and Formula of the Kingdom of Ends) has not been furnished. Second, I extend Hegel's counter-factual contention developed in chapters 1 and 2; namely, that a positing of a different normative specification allows for the opposite surd action to be done – that is, one not only may be untruthful, but according to the Formula of Humanity one *must be* untruthful.

In Section III.B I consider the derivation of duties against suicide and deceit from the *Groundwork*. These derivations differ significantly from the derivations in *The Doctrine of Virtue* in that they rely on a single, common intermediate premise: that the



relevant acts treat the respective agent (oneself or another) as a mere means. However, a closer analysis of why suicide and deception must use an agent as a mere means reveals the susceptibility of these derivations to Hegel's emptiness critique. In the case of suicide, one treats one's own humanity as a mere means through *disregard* – that is, the intentional debasing of one's worth relative to other values. Such disregard, however, is a product of the agent's own intention, and thereby susceptible to Hegel's subjectivism objection. In the case of deception, the agent is used as a mere means insofar as his rational will is circumvented or coopted such that he cannot possibly participate in the end set by the liar. I argue that this criterion is insufficient for showing the impermissibility of an action by showing how numerous mundane actions may satisfy this criterion. As such, therefore independent moral content is necessary to derive a duty against certain classes of deception; I suggest that a candidate for this content is the normative expectations of conventions and social institutions, such as friendship.

Finally, it is important to emphasize that the extension of Hegel's emptiness critique to the Formula of Humanity and Formula of the Kingdom of Ends will not precisely parallel the arguments forwarded against the Formula of Universal Law. Although Kant contends that the various formulations of the Categorical Imperative are essentially equivalent (*G*, 4:435; see also 4:431), there can be no question that they differ in function and application. The Formula of Universal Law, being a test of proposed maxims by means of a universalization procedure, is susceptible to peculiar objections; most notably, that the duty revealed by the universalization procedure (e.g., that false-promising is proscribed) is itself contained in an assumed normative specification, and thus the legislating of the universalization procedure is *superfluous*. No such superfluity

is to be found regarding the Formula of Humanity. Indeed, Hegel has the highest regard for Kant for in placing rational nature at the center of our ethical understanding and considers the moral point of view to be indispensable to rational ethics.<sup>2</sup> Thus, the Formula of Humanity and the Formula of the Kingdom of Ends involve no superfluity; they are merely incomplete and insufficient for providing determinate, concrete duties.

### I. TEXTUAL SUPPORT OF HEGEL'S BROAD CRITIQUE

The emptiness objection is often interpreted as applying exclusively to the Formula Universal Law. This objection is, to a degree, understandable when one considers Hegel's own focus and elaboration of his critique. The emptiness critique is presented in various texts spanning the different periods of Hegel's theoretical development. At no point does Hegel discuss or even appear to acknowledge formulations of the Categorical Imperative in addition to the Formula of Universal Law. In addition to this rather astonishing *absence*, Hegel's positive arguments and characterization of Kant's moral philosophy are often in terms of the Formula of Universal Law. For example, in his *Natural Law* essay, which includes the fullest elaborations of the emptiness critique with concrete examples, Hegel focuses exclusively on the Formula of Universal Law and, more specifically, on the case of keeping a deposit for which there is no proof (see chapter 1). Similarly, in the *Phenomenology* (§§419-429) Hegel follows his critique of 'Reason as lawgiver', where "sound Reason knows immediately what is right and good," with an analysis and critique

---

<sup>2</sup> Regarding Kant's contribution to moral philosophy, Hegel writes: "The essential element of the will for me is duty. . . . I should do my duty for its own sake, and it is in the true sense my own objectivity that I bring to the fulfillment in doing so. In doing my duty, I am with myself and free. The merit and exalted viewpoint of Kant's moral philosophy are that it has emphasized this significance of duty" (§133, *Addition*). For Hegel's view on Morality in relation to ethics, see: Hegel (*PR*, §§106, especially the *Addition*, 112 *Addition*, 141, 149 and 150).

of 'Reason as testing laws'. That is, he seems to move from a discussion of rational intuitionism, where moral laws are known immediately, to a discussion of Kant's moral theory understood exclusively through the lens of the Formula of Universal Law.

And yet, when we look to other sources, or even reconsider these sources carefully, we can see clearly Hegel's intention for the emptiness critique to apply broadly to the moral point of view. A brief but important passage is found in the *Encyclopedia*:

[Kant] expressly recognised the positive infinity of practical reason, specifically by ascribing to *willing* the faculty of determining itself in a universal manner, that is to say, through *thinking*. . . . [B]ut the recognition of this faculty does not yet answer the question of what the *content* of willing or of practical reason is. *Only* with the recognition [of the will's freedom] the question of the *content* of the will or of practical reason is still not answered. When it is said that human beings ought to make the *good* the content of their wills, right away there recurs the question of the content, i.e., the determinacy of content; and we get nowhere either with the mere principle of the agreement of the will with itself or with the demand to do duty for duty's sake. (*EL*, §54A)

Here Hegel clearly has the grounding of duty in practical reason, whatever form that may take, as his critical target and is contending that practical reason alone cannot furnish the needed content for determinate duties. His further contention regarding “the agreement of the will with itself” – which I take to be a reference to the Formula of Universal Law – serves to show that Hegel distinguishes between the general position of grounding morality and duty in practical reason and the more specific means of deriving duty through the universalization procedure and finds *both* to be insufficiently substantive.

The most explicit statement concerning the breadth of Hegel's critical target comes, again, from the crucial §135 in the *Philosophy of Right* – importantly Hegel's most mature work. Here, Hegel presents his critique as generic with regard to *Moralität*, and specifically Kant's moral philosophy, and even distinguishes this generic position from the more specific arguments pertaining to the Formula of Universal Law. He writes:

“to cling on to a merely moral point of view without making the transition to ethics reduces this gain to an *empty formalism*, and moral science to an empty rhetoric of duty for duty's sake.” Hegel continues, contending, as we have seen, that such a point of view may, in order to determine particular duties, bring in “material from outside.” This is the heart of the emptiness objection both as I have explicated it as well as Hegel's Kantian interlocutors. The centrality of this passage to understanding Hegel's critique makes it surprising that Kant's defenders would limit the emptiness objection to the Formula of Universal Law; especially considering that that Hegel continues in this same passage to distinguish this general objection, relevant to *Moralität* as a whole, from the subsidiary arguments applicable to the Formula of Universal Law exclusively: “Kant's further form – the capacity of an action to be envisaged as a *universal* maxim does yield a more *concrete* representation of the situation in question, but it does not itself contain any principle apart from formal identity and that absence of contradiction already referred to” (*PR*, §135). Admittedly, Hegel goes on to clarify *only* this latter, subsidiary contention with the specific cases of property theft and murder, and thus it is again understandable that his interlocutors would engage with these concrete arguments. Nevertheless, however underdeveloped Hegel's arguments may be concerning his more general point, his intention is clear: “no immanent theories of duties is possible” from the moral point of view unless it is augmented with material brought in from outside.

## II. THE FULLNESS OF THE FORMULA OF HUMANITY

Regardless of Hegel's intentions regarding the scope of the emptiness critique, Kantians have good reason in light of the Formula of Humanity to contend that his critique must be

limited to the Formula of Universal Law. Said differently, a consideration of the Formula of Humanity, and perhaps also of the Formula of the Kingdom of Ends, would seem decisively to show that Hegel's general intention must be wrong. This certitude is grounded in the explicit content of the Formula of Humanity: "*So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means.*" (G, 4:429). This formulation is, at least, not formal in the same way the Formula of Universal Law is – that is, the Formula of Humanity is not *procedural*. Rather, the Formula of Humanity posits *humanity* as an end which circumscribes actions as dutiful or contrary to duty. This much, I think, must be conceded: the Formula of Humanity is substantial – or “full” – in that it provides clear moral content in the form of an end of moral action; the pertinent question relevant to Hegel's critique is whether it is sufficiently substantial to derive determinate duties and serve as the foundation to an “immanent doctrine of duties.”<sup>3</sup>

To get a clearer sense of why Kantians think the Formula of Humanity is sufficiently substantial for deriving determinate duties we may look to their interpretation of the Formula of Humanity as *teleological* as opposed to *deontological*. Kant's moral philosophy is rightfully and understandably held to exemplify deontological ethics. Duty is understood solely in terms of actions which conform to and follow from the Categorical Imperative without any consideration of the consequences of the action. The Categorical Imperative “has to do not with the matter of the action and what is to result

---

<sup>3</sup> This means I am defending what Wood calls the “weak thesis.” The “strong thesis” is that the Formula of Humanity, or Formula of Universal Law, cannot provide *any* determinate duties. I think that this is not the case. There are some concrete duties that we can derive from the Formula of Humanity, but they are so limited and restricted that the Formula of Humanity cannot serve as a foundation for a *doctrine* or system of duty.

from it, but with the form and the principle from which the action itself follows; and the essentially good in the action consists in the disposition, let the result be what it may" (*G*, 4:416). This point must be maintained: Kantian morality is unequivocally non-consequentialist; right and duty are always prior to and supersede the good. However, this emphasis on Kantian morality being opposed to consequentialism must not obscure the role of ends, and specifically rational nature as a moral end, in Kant's system.

As Lo (1981) and Wood (1999) both make clear, the role of ends is central to Kant's theory of action, and the *respect and promotion* of rational nature as a moral end is central to his moral theory. First, we may note that Kant holds the capacity to choose and set an end, from among possible alternatives, characteristic of humanity and the property that distinguishes it from animality (*MM*, Intro. Sec. 8; *CBHH*, 8:111-114; Lo, [1981, 184]; Wood [1999, 118-19]; Korsgaard [1996, 110-11]). Moreover, of particular import is Kant's claim that every action must have an end, from which it follows that every duty or moral action likewise has an end:

An **end** is an *object* of free choice, the representation of which determines it to an action. Every action, therefore, has its end. . . . But because this act which determines an end is a practical principle that prescribes the end itself, not the means, it is a categorical imperative of pure practical reason, and therefore an imperative which connects a *concept of duty* with that of an end in general. (*MM*, 6:385)

Kant goes on to argue, by a *reductio*, that among all the ends adopted in free action, some must be the ends of duty, since if all ends were hypothetical "then all ends would hold for practical reason only as a means to other ends." And since "there can be no action without an end, a *categorical* imperative would be impossible" which "would do away with any doctrine of morality" (*MM*, 6:385). Dutiful action, thus, must have an end; and since duty is to act *from* the moral law it follows that the moral law, or in the case of

embodied beings, the Categorical Imperative, must provide an end and one that holds for every rational being. And this is precisely what we see in the *Groundwork* (4:427): “Now what serves the will as the objective ground of its self-determination is an end, and this, if it is given by reason alone, must hold equally for all rational beings.” The content of this unique, objective end we have already seen in the Formula of Humanity; namely, “that the human being and in general every rational being *exists* as an end in itself” (G, 4:428).

The preceding overview has important implications for how we conceive of Kant's moral philosophy. In consideration of these and other passages Wood contends that “if a 'deontological' ethical theory is one that *precludes* grounding a moral principle on substantive value or ends, then the aim of Kant's arguments at G 4:427-432 is to show that no deontological theory is possible” (Wood, 1999, 114). This would of course include Kant's own moral theory and thus implies that Kantianism is not strictly deontological, but teleological.

This conclusion, however, would appear to introduce a tension – or perhaps a straightforward contradiction – into Kant's theory. Lo writes that this teleological interpretation immediately faces at least two objections: First, the teleological interpretation conflicts with Kant's contention “that moral laws have to be categorical” (Lo, 1999, 185), which is to say that they are practically necessary without reference to further ends (see G, 4:414). Said differently, one might argue that an imperative cannot be categorical while simultaneously incorporating an end, since incorporating an end into an imperative renders that imperative hypothetical. So, as Lo writes, is the “conception of 'a teleological categorical imperative' not contradictory in itself?” Second, as we have already seen above, Kant explicitly states that a categorical imperative is not concerned

with “the matter of the action and its presumed results.” As such one might reasonably infer that all categorical imperatives *must* be empty of all content.

Any tension stemming from these objections, however, is resolved through an appeal to Kant's distinction between different kinds of ends. The first telic distinction is between *non-existent* and *existent* ends. As Wood notes, we tend “to think that the concept of an end is *nothing but* the concept of a not yet existing object or state of affairs” (Wood, 1999, 116). But this is not so: an end is, rather, anything *for the sake of which* we act, and this may include already existing objects or states of affairs. Wood gives the example of self-preservation, and I believe we may add more ordinary examples such as marriage or friendship. Importantly, with such ends we do not act to bring about an object or state of affairs, but rather act to preserve or promote the continuance of an object or state of affairs. The second telic distinction is between *subjective* and *objective* ends. A subjective end is the object of inclination – the wants and desires of the individual agent which Kant employs when discussing the various “grounds” of action: a “subjective ground of desire is an *incentive*,” whereas an objective ground of volition is a *motive*. Being the product of inclination, subjective ends are diverse, corresponding to the varied lives and interests of the people who adopt them. Importantly, then, subjective ends do not hold for all rational beings – they are freely adopted or not adopted by the agents themselves. Crucially, if an end is both subjective and conditional (i.e., non-existent), then, Kant argues, the value of the end is relative: “the ends that a rational being proposes at his discretion as *effects* of his actions (material ends) are all only relative; for only their mere relation to a specially constituted faculty of desire. . . gives them worth” (*G*, 4:428). Being determined by reason, objective ends hold equally for all rational



beings. This distinction of ends leads to a distinction between practical principles: a principle is *formal*<sup>4</sup> if it abstracts “from all subjective ends” and *material* “if they put these, and consequently certain incentives, as their basis” (G 4:427). Being determined by reason and holding equally for all rational beings, the worth of an objective end is not relative; that is, the worth of an objective end is not dependent on the adoption of the end by a particular individual with conditional, contingent inclinations.

With these distinctions in mind Kant's theory can be shown to be teleological without contradiction. First, the incorporation of ends *per se* is not inconsistent with categorical principles of action; rather, only the incorporation of ends which are subjective or conditional would preclude categoricity.<sup>5</sup> An objective end serves as the ground of volitional for all rational beings, is thus categorical. Second, Kant is not inconsistent when he says that a categorical imperative is not concerned with “the matter

---

<sup>4</sup> This use must be kept distinguished from Hegel's use; for Hegel, the formalism of the categorical imperative concerns the abstract nature of the right and the good, and consequently duty (§§133-135). For Kant, motivations are formal not in the sense of being abstract, but rather in the sense of not being grounded subjectively, or grounded in the peculiar incentives of the agent. Though distinct, it would be mistaken to say the two senses are entirely unrelated: for Hegel, in order to get content-full duties, formal ends (rational agents) and material ends must be united in ethical life (*Sittlichkeit*).

<sup>5</sup> That a subjective end could not correspond to a categorical imperative seems clear: a categorical imperative holds as valid for all rational beings, but a subjective end holds only for those possessing an incentive to that end. That a non-existent end is equally incompatible with a categorical imperative may not be as evident. Lo clarifies as follows:

Even if the end is not set by our inclination, but by our reason, insofar as it is a producible end, it still cannot be the end of a categorical imperative. This is because when an imperative enjoins an action in order to attain that end, that action has binding force on us only because it is a means to that end.... Once it is no longer an efficient means to bring about that end, it will not be enjoined by the imperative which aims at bringing about that end. Hence its binding force is till conditional and the imperative itself is hypothetical. (Lo, 1981, 186)

If the action will produce the end, then the action is binding on us; however, if circumstances change such that the end will not be realized by the action (e.g., because the end is already realized, or because the end can no longer be realized by these means), then the action will no longer be binding on us. Suppose, for example, that we set the end of maximal welfare, and we determine that respecting one's rights (of property or of bodily security) are an efficient means to this end. If these means are determined to be no longer necessary to the realization of this end, or as no longer an efficient means to the realization of this end, then we will no longer be bound by such rights.

of the action and its presumed results.” The matter of the action refers to those material aspects of the action, which Kant restricts to aspects of the action that have incentives as their basis. In other words, the “matter” of the action does not preclude formal content; that is, content related to rational *motive* (see Lo, 1981, 187; *G*, 4:400 and 427).

Likewise, when Kant speaks of “presumed results” we should restrict his meaning to the *consequences* or *effects* of action; existent ends are not the result or consequence of action, but nonetheless may be promoted, preserved and respected in action.

We already know that Kant holds up *humanity* as an end that is both objective and existent – an end-in-itself – and thus capable of corresponding to a categorical imperative. His arguments for why humanity satisfies these requirements are not, at the moment, our concern. The foregoing has been to establish the substantiality of the Formula of Humanity and the, or at least one, interpretation of Kant that substantiates the claim, which is necessary for showing that the Formula of Humanity may be used to derive determinate duties. However, it must not only be shown that the Formula of Humanity is generally substantial, but further that the abstract content of respecting humanity, in oneself and another as an end and never as a mere means, also produces specific duties. In the following two sections I turn to the various grounds for accepting this further, more specific contention.

### III. THE DERIVATION OF NEGATIVE DUTIES IN *THE DOCTRINE OF VIRTUE* AND THE *GROUNDWORK*

The development of this crucial contention concerning the derivation of particular duties

by Hegel's Kantian interlocutors is a complicated matter. This section addresses the derivation of negative duties first as explicated by Kant in *The Doctrine of Virtue* and, second, as explicated in the *Groundwork* as well as contemporary Kantians who defend and adapt these explications. While the essential framework for the contention is shared among different authors, they diverge in key details. We shall have to explicate and keep these various divergences in mind when applying Hegel's critique.

The first step in understanding the Kantian defense of *Moralität* as a ground for particular duties is to review the derivation of said duties in *The Doctrine of Virtue* in general. This review will provide insight into the general application of the Formula of Humanity and specifically how it functions within a syllogism as some Kantians propose. Following this general review and functional explication we can turn to some more specific and elaborate arguments in defense of specific duties derived from the Formula of Humanity – first as explicated in *The Doctrine of Virtue* and then the *Groundwork*.

Lo provides a helpful enumeration of the particular duties Kant derives from the Formula of Humanity in *The Doctrine of Virtue*. These duties are divided into four kinds: negative (or prohibitory) duties to oneself; negative (or prohibitory) duties to others; positive (or promoting) duties to oneself; and finally, positive (or promoting) duties to others. If we take humanity to be a self-existent end, Lo argues, the negative form of duty immediately follows: “Its existence imposes a 'limiting condition' on our actions. No human being should be 'manipulated' or 'used' as if he were merely a 'thing', but should always be treated as a 'person'” (1981, 190-191). Lo then appears to suggest that we can conclude from this generic negative form of duty derived from the Formula of Humanity itself all of the following prohibitions: with regards to our “animal being, suicide, carnal

self-defilement, self-stupefaction by immoderate use of food and drink are all forbidden; and as a moral being, lying avarice, and false-humility are all prohibited” (191). As for prohibitory duties to others, “pride, calumny, and mockery are all condemned.”<sup>6</sup>

Such are the negative duties enumerated in *The Doctrine of Virtue*; yet our particular duties are not limited to those enumerated. Rather, both Lo and Wood argue that our duties are derived through an application of the Formula of Humanity that takes the form of a syllogism. The structure of this syllogism incorporates the Formula of Humanity itself as the major premise, and then with the introduction of a minor (or as Wood names it, an “intermediate”) premise, we may derive particular moral conclusions (e.g., do not commit suicide, or do not make false-promises).

While there is some agreement concerning the syllogistic form of applying the Formula of Humanity, Lo and Wood diverge when considering the derivation of concrete duties; specifically, with regard to the function of minor premises within this syllogistic deduction.<sup>7</sup> This difference can be presented as two competing theses:

- **Productive Thesis:** The productive thesis is represented by Lo, and holds that the syllogistic derivation of concrete duties is “fecund” and may be achieved without reference to any further moral content.<sup>8</sup>
- **Guiding Thesis:** The guiding thesis – which Wood sometimes appears to defend – holds that the Formula of Humanity is only a “guide,” and that we need to incorporate *logically independent* judgments in the syllogistic derivation of concrete duties.

In defense of the Productive Thesis Lo contends, without elaboration, that the minor

---

<sup>6</sup> Lo enumerates a similar list for positive duties which is addressed in the next chapter.

<sup>7</sup> See also Guyer (2006, chap. 7) who while less specific regarding the form we must use as a supplement to our derivation of duty, does agree with Lo and Wood insofar as we must use our “judgement” to augment the Formula of Humanity.

<sup>8</sup> See, also, Freyenhagen (2012, 53).

premises of the syllogism are “nonmoral” and proceeds to conclude, with the preceding enumeration of particular duties drawn from *The Doctrine of Virtue* yet without any further consideration of Hegel's critique, that “[t]his 'teleological' supreme principle of morality really works, and the fertility, rather than the bareness, of the Kantian Moral Law is revealed” (1981, 195). In other words, he takes his essential work to be done. This is a crucial exegetical contention: If Lo is right, and the intermediary premise is nonmoral, *and* we are able to derive particular duties from the proposed syllogism, then the Formula of Humanity will have been shown to be sufficiently “full.” That is, the Formula of Humanity is the only moral content to which we need appeal in our derivation of particular duties.

Wood's analysis differs both in his characterization of the minor premise and in his unqualified optimism of the “fertility” of the Formula of Humanity itself. Wood's position is sometimes ambiguous: On the one hand, Wood does not argue that the intermediary premise is “nonmoral,” but rather merely that this premise must be “logically independent from FH” (Wood, 1999, 152). By “logically independent” Wood means that we can deny the minor premise while accepting the major premise – that is, the Formula of Humanity – which of course implies a rejection of the particular moral conclusion. On the other hand, Wood contends that the intermediate premises, while contentious, are empirical in nature and any contentiousness may be resolved through progressive enlightenment. It is important to emphasize that the minor premise is absolutely essential to the derivation of the particular moral conclusion. Therefore, if this is the case and the intermediate premise can be shown to either contain moral content or for its acceptance to require independent moral judgment, then we will have shown the

formal emptiness of the Formula of Humanity. In other words, by characterizing the minor premise as “logically independent” Wood has not eliminated the possibility that the minor premise itself contains moral content or requires independent moral judgment. This should raise a critical and distinctively Hegelian red flag: is it the case that the Formula of Humanity cannot function to provide particular duties without “importing content from outside” in the form of independent moral judgements?<sup>9</sup>

---

<sup>9</sup> I hasten to add that Wood, for his part, perceives the disputability of the intermediate premises and the challenge this poses for the Kantian position. Unlike Kant – and Lo – Wood does not possess certainty regarding the derivation of specific duties, especially when these duties are construed in absolute terms. Indeed, Wood’s view approaches what I would call a Hegelian position, save for two crucial points: first, he characterizes the controversial intermediate premises as *empirical*; and two, he appears to still endorse a trans-cultural, trans-historical, trans-discursive standpoint for justifying reasons given in support of intermediate premises. Regarding the challenge facing the derivation of duties, Wood writes:

But the use of FH in moral deliberation always requires an intermediate premise...and such premises are often doubtful or controversial. We badly misunderstand Kant’s theory if we suppose he thought the *a priori* principle of morality (in any formulation) could determine what to do apart from such empirical principles. (1999, 154)

Characterizing the intermediate premises as empirical perhaps explains his optimism in the succeeding paragraphs:

The intermediate premises are disputable in character because rational nature reveals itself only under particular cultural and historical circumstances and our views on it are corrigible (as we learn more about ourselves) . . . We do have quite a bit of knowledge about rational nature, however.... This knowledge is what we must use to guide our judgements about how the dignity of humanity should be respected in action.

Wood then goes on to warn that some of what counts as knowledge about rational nature is actually “pernicious prejudice, which it is the task of enlightenment to remove” (*ibid.*) Thus, Wood is optimistic about gaining knowledge about what morality requires, and specifically the Formula of Humanity, because he is optimistic about scientific, philosophical, and cultural progress concerning our knowledge of human nature (for a similar view, see Herman [1993e, 88-91]).

This view is to a limited degree compelling: the knowledge we have gained about races and the sexes has lead us to abandon the relevant prejudices held by Kant; it is also true that we will undoubtedly learn about our nature, including our rational nature, as a species and this may continue to shape our application of the Formula of Humanity. However, what this position ignores, and what I aim to show in the following critique, is that *moral* judgments and not merely empirical judgments are crucial to the derivation of duties; moreover, disputes over such judgments are not resolved through increased scientific, philosophical, and cultural knowledge of rational nature. This latter point is illustrated, briefly, by Wood’s own thoroughly unsubstantiated claim that the false-promising intermediate premise is “easier to accept” than the suicide intermediate premise because “the direct frustration of a rational being’s agency through deception is by virtually all accounts a more direct affront to the worth humanity than a rational being’s termination of its own life process” (*ibid.*). Thus, Wood in defending Kant – or an updated Kantian position – has concluded that deception is more clearly impermissible than suicide because most people (which people in what time and place is not mentioned) are more *offended* by deception. And further, we are led to believe that such near

Interestingly, as we shall see below, Kant's own derivation of particular negative duties from the Formula of Humanity utilizes ancillary syllogism and moral judgments independent of the Formula of Humanity. Specifically, Kant appeals to the illicitness of acting contrary to the natural purposiveness of language and to the supposedly intrinsic debasement of humanity involved in suicide. Ironically, then, we find resources for rebutting the strong claims of Kant's defenders regarding the fecundity of the Formula of Humanity in Kant's own explication.

We must keep in mind that when applying Hegel's emptiness critique to the Formula of Humanity both theses must be addressed. As Lo does not provide any elaboration or formalized examples of "nonmoral" minor premises, I shall not argue directly against the Productive Thesis. Rather, I shall provide examples of syllogisms reconstructed from Kant's own arguments and from the examples provided by Wood. In considering these examples, my goal is to identify the intermediate premises as not only logically independent, but also to show how in every case of a derived particular duty, this derivation relies on a logically independent content that is both moral and controversial, and thus in need of independent *justification*.

---

unanimity of offense is the product of a progressive enlightenment. And yet in evident contrast to this position which explicitly relies on mere agreement as to what *offends*, Wood still insists that "such intermediate premises may be accepted or rejected on the basis of reasons; they do not depend merely on blind intuitions or irrational prejudices" (55). In what follows I aim to both contest the intuition that deception is what accounts for our rejection of lying; but more importantly, show that reasons may historically, socially, and conventionally contingent and these contingent reasons and their essential role in determining duty threaten the Kantian project.

### A. The Derivation of Negative Duties in *The Doctrine of Virtue*

As enumerated above, Kant categorizes duties as either negative (prohibitory) or positive (promoting), and further divides these into duties to oneself and duties to others. I will limit my discussion to two negative duties: the duty not to commit suicide and the duty not to lie, which represent duties to oneself *qua* animal being and duties to oneself *qua* moral being respectively. The critique presented of these negative duties is meant as a heuristic; the same critique may be applied in, with various appropriate modifications, to other negative duties derived from the Formula of Humanity.

#### 1. *Suicide*

In *Book I of Doctrine of the elements* Kant considers negative duties to oneself, the first of these being a duty opposed to self-killing. This duty derives, according to Kant, from the more general duty of a human being to “preserve himself in his animal nature.” It is because, Kant says, we are considering negative duties that we must consider “the *vices* opposed to duties to oneself” such as willful physical death. Further, in these passages Kant is not only considering killing oneself understood as *suicide*, which he characterizes as the total killing of oneself, but also *self-mutilation*, which he characterizes as the partial killing of oneself.<sup>10</sup> A final prefatory note, Kant here is considering the prohibition against suicide as a violation of the duty one has to oneself *qua* animal being. He recognizes that suicide (or self-mutilation) may also be considered as a violation of one’s duty to others (e.g., one’s spouse or children), to one’s fellow citizens, or even to God.

---

<sup>10</sup> Kant further distinguishes between material and formal self-mutilation. The former being physical maiming, while the latter is the deprivation of “one’s capacity for the natural *use* of one’s powers” (*MM*, 6:421).



These qualifications must be kept in mind when considering both the arguments for the proscription as well as the subsequent critique.

Turning first to suicide properly speaking, Kant argues that such willful physical death violates our duty to our self *qua* animal being since,

[t]o annihilate the subject of morality in one's own person is to root out the existence of morality itself from the world...even though morality is an end to itself. Consequently, disposing of oneself as a mere means to some discretionary end is debasing humanity in one's person to which the human being was nevertheless entrusted for preservation. (*MM*, 6:423)

Wood (1999, 152) reconstructs the argument as a syllogism with the following intermediate premise:

1. **FH:** Always treat humanity, in your own person and in the person of another, as an end-in-itself and never as a mere means.
2. **Ps:** The act of suicide always fails to respect humanity in one's own person as an end in itself.
3. **Cs:** Therefore do not commit suicide.

We may note immediately the logical independence of Ps and FH. Per Wood's explication, we can accept FH and reject Ps. The question is whether this independent premise also contains distinct moral content; a question we must answer in the affirmative. The intermediate premise makes a distinct claim; namely, that suicide fails to respect humanity and thereby constitutes a violation of the Formula of Humanity; such a claim represents a distinct moral judgment. As such, independent justification is needed for Ps, and such independent moral justification consists in material that must be brought in from outside. On our initial assessment, then, we can already conclude that the Formula of Humanity is not itself sufficiently substantial for deriving particular duties: we must make logically independent moral judgements regarding this intermediate premise. (This point will recur with our consideration of other duties.)

It may seem that the need for independent justification is clear from Kant's own explication (as well as Wood's assessment of the argument, as we will see below) in that Kant appears to provide his own justification for the intermediate premise: Ps is supported by the claim that suicide "roots out" the existence of morality, and thus "debases" humanity, and it is only as a consequence of these facts that suicide is proscribed. Thus, in order to derive the proscription against suicide, Kant actually provides an ancillary argument, which reconstructed, takes the following form:

1. **Pas1:** "To annihilate the subject of morality in one's own person is to root out the existence of morality itself from the world, as far as one can" (6:423).
2. **Pas2:** To root out the existence of morality from the world, as far as one can, always debases morality itself.
3. **Cas:** "Consequently, disposing of oneself as a mere means to some discretionary end is debasing humanity in one's own person" (6:423).

Some observations: First, I take the conclusion of this argument (Cas) to be equivalent to Ps above: to debase humanity is to fail to treat humanity as an end in itself, since it consists in lowering the worth of humanity as an end to the same order as other, discretionary ends. Second, Pas2 is absolute, allowing for no exceptions to the proscription. Third, the argument seemingly moves from absolute terms to conditional terms – the argument concludes that suicide is wrong only if it is used as means to "some discretionary end." By "discretionary end," however, I take Kant to mean any end that is the product of subjective choice, and thus any end that is not an end-in-itself. This may appear to return the argument to absolute terms; however, as I will discuss more below, it may leave room for suicide when the act is for another end-in-itself – that is, another person.

Does this argument represent independent justification for Ps? Perhaps not, as this argument may be interpreted such that no material from outside need be brought in to

justify Ps: If suicide just is – analytically – the rooting out of morality from the world (as far as one can), *and* such an act necessarily involves or consists in debasing humanity, then Kant has not appealed to an independent moral justification, but only to an analysis of the act itself. Thus, our strategy must be to question this analysis.

Turning to the premises, then, for now Pas2 must stand: if suicide does in fact consist in the rooting out of all morality, it is hard to see how this would not consist in a debasement of morality. However, Pas1 is immediately dubious and in need of modification. Pas1 states that suicide roots out “the existence of morality from the world”; but this must be recognized as a separate (and, morally speaking, more severe) effect than that brought about by the mere killing of oneself. In other words, Kant here is certainly wrong, or is at least imprecise, concerning the consequences of suicide. In killing oneself one is not annihilating the source of morality *completely*; which is to say, the rationally free will he calls “humanity.” Indeed, by the time *The Doctrine of Virtue* was written Kant had already laid the foundation for his arguments, including already establishing (or, rather, taking himself to have established) that the source of morality and the object of respect I find in myself is necessarily shared by *all* rational beings (G 6:428-429). Thus morality, the ground of morality, and the object of respect remains in others. Morality, per se, is not annihilated through suicide. Rather, in committing suicide one merely roots out morality as found in oneself. Thus, we should reject Pas1, and correlatively Pas2, and replace them premises that more accurately represent the consequence of suicide:<sup>11</sup>

1. **Pas1\*:** To annihilate the subject of morality in one’s own person is to root out the

---

<sup>11</sup>This rejection is supported by Kant's own defense of Cato, considered below. If suicide really were the rooting out of all morality as far as one can, then Cato's action would not furnish an example of a virtuous suicide.

- existence of morality as found in oneself.
2. **Pas2\*:** To root out the existence of morality, as found in oneself, always debases morality itself.
  3. **Cas:** “Consequently, disposing of oneself as a mere means to some discretionary end is debasing humanity in one’s own person” (6:423).

With these modified premises we can turn to evaluating Kant’s argument, which will involve a renewed consideration of the intermediate premise.

The argument is valid. The question is whether we must accept the premises. Specifically, is it always or necessarily true that the annihilation of the moral subject debases humanity? Wood, for his part, questions the absolute terms of Ps – and by extension Pas2\* – and recognizes that many will reject the proscription against suicide under certain conditions:

For example, some people think that it is degrading to humanity for a person afflicted with a horrid and debilitating terminal disease to live on in a virtually subhuman condition... [and] two people who equally accept FH can therefore disagree about a conclusion such as Cs because they disagree about Ps. (Wood, 1999, 152-153)

Interestingly, Wood also notes Kant’s own sympathy with such conditionality of the proscription against suicide. Wood cites a passage from Kant’s *Lectures on ethics* which is worth quoting in full:

Cato killed himself when he saw that it was not possible to evade the hands of Caesar, though the whole people still depended on him; as soon as he had subdued him, as the defender of freedom, the others would have thought: if Cato submits, what shall we do? But if he killed himself, then the Romans could devote their last ounce of strength to defending their freedom; so what should he do? It appears, then, that he saw his death as necessary; he thought: since you can no longer live as Cato, you cannot go living at all. In this example, one must freely concede that in such a case, where suicide is a virtue, it has a great appearance of plausibility in its favor. But this is also the sole example that the world furnishes us in which suicide can be defended. (*LE*, 27:370-371; quoted in Wood, 1999, 153 n. 37)

In this passage Kant himself does not fully condone a conditional proscription against

suicide; but, to the contrary, contends that Cato's predicament and virtuous suicide is unique in history. Wood rightly rejects this latter claim as extremely implausible, and suggests that many of us – and specifically those facing a dehumanizing terminal illness – may make a similar, virtuous declaration: that if we cannot live as ourselves, we cannot go on living at all. If Kantians concede this point, then there are situations, contexts and objectives that do furnish good reasons for committing suicide. What the passages leave unclear is the means by which we identify these good reasons.

This latter point is worth emphasizing: both Kant and Wood fail to appreciate what these exceptional cases represent. That there are exceptional cases is a problem; but the deeper problem lies in the fact that neither Kant nor Wood has identified the grounds by which disagreement over which actions do or do not respect humanity may be resolved. Put differently, both Kant and Wood recognize that there may be good reasons for committing suicide. Abstractly, both would insist that such good reasons must be in accordance with the Formula of Humanity – that is, the act of suicide must not debase humanity. But this fails to answer the immanent question of when, and under what conditions, suicide debases humanity.<sup>12</sup>

To put this challenge in concrete terms we may turn to any number of cases. Most

---

<sup>12</sup> One might respond that criteria have been given: for Kant it is the defense of freedom; for Wood, it is the preservation of "living as oneself." If it is the preservation of freedom, then what sort of freedom is being preserved? If it is humanity, then Cato appears to be treating the worth of his humanity quantitatively – by trading his humanity for others – and is violating Kant's insistence that one has the duty not to commit suicide as long as one lives. Maybe the thinking is that Cato is going to die anyway, so he is not destroying anything that won't be destroyed soon, and is, moreover, saving the humanity of others. But is this not true of all of us? Are not all of us mortal and soon to die? Does it follow, then, that may we commit suicide for other's freedom whenever it is efficacious? Perhaps, by "freedom" Kant means "liberty" in this context. If this is the case, then liberty is seen by Kant as a good for which humanity – in oneself at least – may be sacrificed. As for Wood's apparent suggestion, that we may commit suicide when and if we will no longer "be ourselves," this raises the question of what it means to "be ourselves." That is, what might count – across culture and history and discourses – as not being our self?

prominent are Kant's own casuistical questions. Immediately following his primary exposition Kant considers several difficult cases where suicide may be permissible: these include sacrificing oneself to save one's country, Frederick the Great's willingness to poison himself to avoid "ransom harmful to his state" in case of capture, and a man afflicted with hydrophobia who wishes to avoid harm to himself or others (*MM*, 6:423-424). Kant presents these casuistical cases as difficult exceptions; however, it is easy enough to expand on such cases with examples both historical and contemporary. Contra Kant's claim of an isolated exception for Cato, we may cite various historical cases: for example, the suicide of Hirate Nakatsukasa Kiyohide who committed *kanshi* (a form of *seppuku* performed as a means of remonstrance) in order to persuade his lord, Oda Nobunaga, to abandon his disgraceful behavior. In more recent history, we may consider the suicide of Thích Quảng Đức who protested the abuses of the South Vietnamese government by performing self-immolation. There are countless additional cases from the context of war: historically, Texans celebrate and honor the heroes of the Alamo who faced certain death in order to stymie the advancing Mexican army, but of course acts of equal valor continue. Some examples may accord with our cultural values or intuitions, such as the case of a mother sacrificing herself to save her son.<sup>13</sup> But, importantly, other cases are ostensibly contrary to our current cultural values and intuitions and still others challenge the limit of these values and intuitions. An example of the former would be the

---

<sup>13</sup> Such cases, though the continued and too often need for them is deeply regretful, impugn rigorism concerning suicide: "Among the 49 victims of the Sunday's Orlando massacre was Brenda Lee Marquez McCool. McCool, 49, was killed by shooter Omar Mateen when she shielded her 21-year-old son, Isaiah Henderson. She was not only a mother of 11, but also a two-time cancer survivor." *Orlando mom of 11 sacrificed herself to save son in shooting* [Online]. Available: <http://www.aol.com/article/2016/06/14/orlando-mom-of-11-sacrificed-herself-to-save-son-in-shooting/21395172/> (accessed March 23, 2017).

Eskimo practice of senicide (see Leighton and Hughes, 1955); the latter is exemplified by the controversial issue of voluntary euthanasia for healthy individuals (see Appel, 2009; see also Ost and Mullock, 2011). Such cases that are contrary to or challenging of our current intuitions are of special import because they represent the cultural, historical, and discursive variability of the scope of contentious cases.<sup>14</sup> Contrary to Cato's suicide representing an isolated example of virtuous self-killing, *Moralität* is faced with diverse and varied cases of self-killing, each of which provides a different possible reason for suicide and thus uniquely challenges the absolute prohibition forwarded by Kant.

Importantly, one cannot claim that each of these above cases furnish a good reason for suicide, but this agnosticism illustrates the deeper point: to identify which cases provide good reasons for committing suicide would require identifying the ground or grounds for determining whether or not a given instance of suicide debases humanity. That such a ground is absent constitutes Hegel's charge of emptiness. Hegel's claim, to recapitulate, is: first, no justification for Pas2\* has been furnished; second, a defense of Pas2\* is necessary for the tenability of Ps, and thus the soundness of the syllogistic derivation of a proscription against suicide; and, third, if a good reason for or against suicide were furnished, this would not justify a general or *absolute* proscription against suicide; and, finally, such a reason would come from outside the Formula of Humanity.

This claim should neither be over- or understated. The above critique of the

---

<sup>14</sup> A Kantian may argue, as I will consider below, that there are exceptional cases, but the scope of these exceptional cases – how many cases are in the undecided penumbra – is narrowing through argumentative dialectic. Though there may be growing consensus regarding controversial cases (though there may not be), unless the rational foundation for such consensus can be identified, the specter of pluralism cannot be avoided. It should also be noted that some of these cases (the heroes of the Alamo and Eskimo senicide) may be permissible on the basis of an absence of an intention to die – the Texans, after all, wanted to *successfully* defend the Alamo and would have welcomed a defeat of the Mexican army. I consider the implications of this form of solution below (see 195ff).

sylogistic derivation of a proscription against suicide should not be taken as concluding that there is no duty not to commit suicide. On the contrary, if we accept the Formula of Humanity, then suicide without good reason (e.g., a nihilist who commits suicide “for no reason at all” or, perhaps, a heartbroken teen who commits suicide capriciously after a breakup) as well as suicide done specifically out of self-hatred or self-debasement will be straightforwardly proscribed.<sup>15</sup> However, this concession is not at odds with the objective of Hegel’s critique, which is to show that no immanent doctrine of duties is derivable from the Formula of Humanity. What an extension of Hegel’s critique has accomplished is a virtual inversion of what is general and what is exceptional vis-à-vis the proscription against suicide: rather than a general proscription against suicide, what is tenable is a proscription against suicide circumscribed to explicit violations of the Formula of Humanity or acts of suicide that are de-contextualized from any justification given for the act – thus, suicide is wrong for the anonymous, self-loathing or morally puerile subject *only*. As soon as we contextualize and particularize the subject, the wrongness of suicide depends on how we judge the moral relevance of the social context and particularity in question. This judgment is, according to Hegel, a matter of *Sittlichkeit*.

## 2. *Lying*

A comparable critique may be given for the derivation of a proscription against lying. This critique, however, is complicated by Kant’s own explication which offers conflicting justificatory premises.

---

<sup>15</sup> In other words, I concede and a Hegelian may concede, that a premise like the following is tenable:

**Pas2<sup>†</sup>:** To root out the existence of morality, as found in oneself, either without cause or reason, or out of self-hatred/debasement, always debases morality itself.



Like suicide, Kant in this section of *The Doctrine of Virtue* is considering lying *qua* violation of one's duty to oneself; though, unlike suicide, Kant considers the duty of rectitude<sup>16</sup> to be toward oneself *qua* moral being, and not animal being. Thus, this duty is contrasted with the duty of rectitude one has to others, and the ethical wrong of lying in the context of duty to oneself is contrasted with the harm done to others when one lies. Moreover, the ethical wrongness of lying is also contrasted with the possible imprudence of lying – for, Kant contends, if lying were a mere violation of a “pragmatic maxim...it could not be considered a violation of duty at all” (*MM*, 6:429).

So if a lie is neither wrong in virtue of the harm done to others, nor the prudential harm done to oneself, where does the wrongness lie? Kant locates the wrongness of lying in the egregious debasement of one's own personality: “But communication of one's thoughts to someone through words that yet (intentionally) contain the contrary of what the speaker thinks on the subject is an end that is directly opposed to the natural purposiveness of the speaker's capacity to communicate his thoughts, and is thus a renunciation by the speaker of his personality.” This, according to Kant, is such a grave infraction of the Formula of Humanity that the lying agent “has even less worth than if he were a mere thing” insofar as a thing “as something real and given” is “serviceable” whereas a liar is not. It is the “greatest violation” of duty to oneself *qua* moral being a person can commit (*MM*, 6:429).<sup>17</sup>

It is important for the succeeding discussion to note the scope of Kant's

---

<sup>16</sup> Kant distinguishes between “honesty” – which is truthfulness in one's declarations – and “sincerity” – which is truthfulness in one's promises. “Rectitude” encompasses both honesty and sincerity (*MM* 6:429).

<sup>17</sup> Even if we agree that untruthfulness, under various circumstances, is wrong, many will reject as outrageous this claim regarding the degree to which lying is wrong. Wood (2008) attributes this extreme view to a pedagogic hyberbolism (see chapter 3).

condemnation of lying: As already noted lying need not be harmful to others or oneself in order to be contrary to duty. Beyond these considerations, Kant also insists that lying is still wrong – and greatly contemptible – even if “done merely out of frivolity or even good nature; the speaker may even intend to achieve a really good end by it. But his way of pursuing this end is, by its mere form, a crime of a human being against his own person and a worthlessness that must make him contemptible in his own eyes” (*MM*, 6:430).

We may again reconstruct Kant's arguments in syllogistic form<sup>18</sup>:

1. **FH:** Always treat humanity, in your own person and in the person of another, as an end-in-itself and never as a mere means.
2. **PI:** In lying a person renounces his personality, and in so doing fails to treat humanity (in his own person) as an end.
3. **CI:** Consequently, one ought not to lie.

As in the reconstruction of the argument prohibiting suicide, the intermediate premise (PI) is both necessary for the derivation of the proscription against lying and logically independent of the Formula of Humanity. Moreover, it makes a moral judgement – namely, that lying constitutes a renouncing of one's personality and thereby fails to satisfy the Formula of Humanity. As such PI is in need of independent *moral* justification. Again, as in the case of suicide, this need for additional justification is clearly seen in Kant's own exposition: the independent justification proposed by Kant is that by lying a person acts in opposition to the natural purposiveness of language and in doing so renounces his personality and annihilates his dignity *qua* human being. Thus, we can

---

<sup>18</sup> Wood's reconstruction of Kant's argument, and specifically his formulation of the relevant intermediate premise, does not follow Kant's justification of the proscription against lying in the *Doctrine of Virtue*, but rather follows the derivation of the proscription against lying as found in the *Groundwork*. As such, I will first consider Kant's own arguments and subsequent discussion and then turn to (a rather lengthy) discussion of Wood's position, which is shared by Korsgaard in Section III.B.

construct an ancillary syllogism for the derivation of the proscription against lying as follows:

1. **Pal1:** When an agent lies, he acts against the natural purposiveness of language, which is the capacity to communicate his thoughts.
2. **Pal2:** Acting against the natural purposiveness of one's capacities constitutes a renunciation of one's personality and the annihilation of one's dignity as a human being. (6:429)
3. **Cal:** Therefore, by lying a person renounces his personality.

In this reconstruction of Kant's ancillary justification for the proscription against lying, both the first and second premises are controversial. Regarding Pal1, no justification or evidence has been provided to support the claim that communicating one's interior thoughts is the natural purpose of human language. This, however, is an empirical question that I will set aside. Regarding Pal2, Kant has not provided any justification for the contention that opposing the natural purpose of a capacity constitutes the renunciation of one's personality.

On this point, Kant appears to be relying his earlier analysis concerning the "objective division of duties" (*MM*, 6:419; see also Aune, 1979, 180) where he distinguishes between formal and material duties one has to himself, which correlate to negative and positive duties, respectively. Regarding both formal and material duties, Kant argues that we may understand our duties in terms of our nature: first, one has a duty to *preserve* his nature and "moral health"; this duty may be summarized as "live in conformity with nature." This duty corresponds to our negative duties, such as a duty not to lie. Second, we have a duty to perfect our nature which corresponds to our moral *prosperity*; this corresponds to our positive duties of natural and moral perfection. With regards to lying, Kant's argument seems to be that by renouncing one's linguistic nature – that is, the capacity to communicate one's ideas and thoughts, etc. – one compromises an

essential capacity for realizing one's rational nature. Simply, one cannot as effectively pursue ends set by reason if one cannot use language as an effective means. This is why we are renouncing our personality – our moral selves; we are, in effect, undermining our rational nature by forsaking our linguistic nature.

Kant's analysis assumes that acting against the natural purpose of language (conceding, for now, this empirical claim) necessarily involves a renunciation of our personality or entails undermining our natural capacity for realizing our rational nature. This is an assumption that may be rejected. Language, and other natural capacities, may be appropriated and adapted in ways contrary to their natural purpose that, nonetheless, contribute to our rational ends. (Falsehoods, for example, may be used as a means of building solidarity our social relationships that, in turn, further our individual or mutual ends.) Kant and Kantians, it would seem, must adopt a modified premise for this ancillary argument, such as the following: **Pal2\***: *Modifying or adapting a natural capacity contrary to the effective realization of our rational ends constitutes a renunciation of one's personality and the annihilation of one's dignity as a human being.* Such a qualified premise, however, cannot justify an absolute but merely socially, historically, and culturally contingent proscription against lying. That is, the social context will be the co-foundation of the duty since the social context will determine how natural capacities are used effectively.

All of the above explication and critique is complicated by Kant's discussion of lying, and the wrongness of lying, in his *Casuistical questions*. There, Kant writes of at least two possible exceptions – or, rather, cases – where latitude may be appropriate. The first is the polite valediction in a letter, where one (following a now antiquated custom)

signs his name “your obedient servant.” In this case, Kant writes, “no one is deceived” and thus the wrongfulness of lying is put into question. In the second case Kant considers whether one may respond politely when asked what one thinks of another’s work – the presumption being that the agent feels tempted to lie as they do not think very highly of the work. May one follow convention in such cases?

These questions are notable for a variety of reasons. First, and perhaps most significantly, the casuistical questions are noteworthy because the criterion for why lying is wrong has changed entirely.<sup>19</sup> The criterion given in Kant’s exposition, as we’ve just seen, is that lying – which is construed here as the intentional use of words contrary to what the speaker thinks – opposes the natural purposiveness of language which is the capacity to communicate one’s thoughts (*MM*, 6:429). Thus, if we remain true to this criterion, then both casuistical cases are clearly, unavoidably wrong: in both cases the speaker is intentionally using words contrary to what he thinks – that is, the speaker is strictly *untruthful*. But in the first casuistical case this rigorous standard of duty is abandoned for the more lenient criterion of whether or not someone is actually *deceived*.<sup>20</sup> However, this justification is problematic for two reasons: First, Kant appears

---

<sup>19</sup> This point is worth emphasizing: The casuistical questions are raised as possible exceptions and thus as counter-examples to the accusation that Kant is overly rigorous. However, this view would imply that the casuistical cases serve as exceptions to the standard justification for the proscription. What we in fact find are new and different justifications being given for the proscription: deception replaces (or is added to) acting contrary to the natural purposiveness of language, and propriety supersedes deception.

<sup>20</sup> There is some controversy in the Kantian literature regarding what Kant means to proscribe by means of Formula of Humanity. Specifically, James Mahon (2009) argues that Kant’s intended proscriptive target is *deception* and not *untruthfulness* per se. This passage challenges this interpretation, first, because unclear why Kant would distinguish between someone being and not being deceived if deception was the original target of his proscription; and, second, because in the second casuistical case, deception is the intended result, and yet the permissibility of the action remains in question. Korsgaard, in her discussion of the proscription of lying drawn from the *Groundwork*, also limits the duty not to lie to a duty not to deceive, comparing this violation of duty to one of coercion. I address these views below in Sub-section B.

to abandon this more lenient criterion as well. In second the casuistical case, deception for purpose of protecting another from insult is the intended and permissible effect of the proposed false-hood. That is, the question presented is whether deception itself is sometimes appropriate. A second, more formally difficult problem is that if we were to apply the justification given in the first casuistical case generally, it would be circular. As an example, consider again the second casuistical case: In this case, the maxim in question is whether or not the agent can tell a false-hood in order to successfully communicate the opposite of what he thinks in order to avoid intentionally revealing a hurtful truth. From this we see that the intention is to deceive, such that the maxim may be reconstructed thus: “May I *deceive* another to avoid a hurtful truth?” If this is an accurate reconstruction of the maxim, then the justification for the proscription is clearly circular: the maxim to *deceive* is wrong because it is *deceptive*. But this provides no justification. What we need, rather, is a justification for why deception is wrong.<sup>21</sup>

The casuistical questions are of further interest because Kant not only abandons the criterion for what constitutes acting against the *natural purposiveness* of language (i.e., abandons untruthfulness for deception), he also apparently abandons natural purposiveness as the justification for why lying is so contemptable and, by appealing to the social expectations of answers and the “politeness” of valedictions, introduces conventional purposes as a replacement criterion. If we follow Kant down this justificatory labyrinth we may again see the relevance of Hegel’s counter-factual contention that was introduced and developed in chapters 1 and 2. If our concern, like

---

<sup>21</sup> Those who derive the proscription not to lie from the *Groundwork* provide such a justification; namely, that lying treats the agent as a means to a discretionary end. I consider this argument in full below in Sub-section B.

Hegel, is surd *actions*, then conventional norms will determine when and under what conditions one is deceived; that is, whether a false-hood is a deception. Consider again the counter-factual normative specification of *Honor-promising* (i.e. the counter-factual convention where in accordance with normative expectations a person may give and may be required to give a false-promise in order to protect his honor, but also the honor of the lender). In this convention, the norm is such that to expect a demand repayment would be disgraceful, but so is asking for money without offering to repay. The consequence is that one may not honestly promise to repay, in doing so he would violate the norms of the convention and, if universalized, no one would accept the genuine promise. Therefore, within a competing normative convention a false-hood may be what is morally required in order to avoid deception and to ensure that one respects another as an end. These considerations militate in favor of the same conclusion reached chapter 2: namely, that any surd action *whatsoever* concerning truthfulness or untruthfulness may be permissible or even what is required by duty given the right normative specification. But such a relativism with regard to particular duties (though not yet with regard to abstract duty) is directly contrary the primary justification of the proscription given by Kant.

With the preceding critical reflections we have not yet considered the intuitive case against the rigorism of P1 – that there are no exceptions to the contemptibility of lying and hence the proscription against it. Most notable is Kant's infamous insistence that lying to the murderer at the door is impermissible. Many find this extreme stringency sufficient for repudiating Kant's position (though admittedly some argue that Kant's refusal to compromise is a sign of personal and not of theoretical strictness). I have addressed this particular case in great detail in chapter 3. For our present purposes,,

however, the point is that this case should again not be considered an exception to an otherwise practicable moral principle: first, we have not been given the justification for the derivation of the *general* proscription; and second, there are many instances in which lying, understood as untruthfulness, may be permissible or even appropriate. However, since these counter-examples serve in a more direct role in impugning the derivation of the duty not to lie as found in the *Groundwork*, in order to avoid needless repetition I will postpone considering them until the next section.

Thus, as it stands, in Kant's own text and derivation of duties the casuistical cases do not reveal the mere possibility of some latitude, but rather present a confusing and ambiguous justification in support of the intermediate premise. This contrasts with Wood's point that the casuistical questions raise possible exceptions for extraordinary cases. The latter interpretation would imply that an exception might need to be made to the generally applicable justification for the proscription. However, this is not evident from Kant's exposition; what we in fact see are new and different justifications being given for the proscription against lying, as well as the absence of these proscriptive justifications (in the case of deception) justifying the exceptions. In conclusion, as in the case of suicide, Hegel's contention has been satisfied: first, no consistent and univocal justification for Pal2 has been furnished; second, Pal2 itself has been called into question; and as the intermediate premise of the ancillary argument in support of P1, the soundness of the syllogistic derivation of a proscription against lying presented in *The Doctrine of Virtue* has not been established. Further, the derivation also faces the challenge of accounting for strongly intuitive counter-examples and justifying an apparently integral and inexorable rigorism. Crucially, however, the most important conclusion for our



purposes is that any rational resolution to any of the above problems will require an appeal to logically and morally independent content, and thus, concerning the derivation of negative duties in *The Doctrine of Virtue*, the Formula of Humanity is neither full nor fertile.

### **B. The Derivation of Negative Duties in the *Groundwork***

A different approach to the derivation of negative duties focuses on Kant's *Groundwork* and contends that negative duties can be derived from the Formula of Humanity directly without appealing to secondary principles. This derivation contends that the prohibited acts in question (I will again limit my discussion to suicide and lying) straightforwardly treat humanity, either in one's own person or another's, as a means to some discretionary end, and thus violate the Formula of Humanity. These derivations remain syllogistic; the syllogisms in the *Groundwork* differ from those in *The Doctrine of Virtue* in that for every derivation of a negative duty, the major premise is the Formula of Humanity and the minor premise consists in the claim that the act in question uses the relevant agent (either oneself, as in the case of suicide, or another, as in the case of lying) as a mere means.<sup>22</sup> Accordingly, I adapt Hegel's critique to this more direct approach.

This adaptation has two strategies: First, I argue that in the case of suicide the action itself cannot be judged a violation of the Formula of Humanity unless specific intentions are attributed to the agent. However, the intentions adopted in an act are determined by the subject alone. Thus, this form of Hegel's critique parallels the

---

<sup>22</sup> It is worth noting that though each derivation integrates as a minor premise the claim that the relevant action uses the agent as a mere means, the sense of "use" changes from argument to argument. In the case of suicide, the sense of "use" is one of disregard; in the case of lying, the sense of "use" is one of manipulation or circumvention.

argument against “natural actions” developed in chapter 2; namely, that the judgement of an act as permissible or impermissible relies on the subjective determination of the intention of the action. This introduces, according to Hegel, an unacceptable subjectivism to our moral judgements and has dire consequences for our ability to adjudicate between permissible and impermissible actions. Second, I argue that in the case of lying – which is a conventional action – adequate criteria for distinguishing between permissible and impermissible acts have not been furnished. In other words, *deception* is not itself sufficient for showing that the Formula of Humanity has been violated; rather, further criteria must be appealed to in order to show why and how deception treats another as a mere means. This criterion, however, is both excessive, in that it leads to the condemnation of benign actions, and on reflection is irrelevant to our actual judgements of what is right and what is wrong in light of the Formula of Humanity. In both cases, then, the Formula of Humanity is shown to be insufficient on its own to derive negative duties.

### 1. *Suicide*

Kant's own argument in the *Groundwork* for a proscription against suicide is brief:

If he destroys himself in order to escape from a trying condition he makes use of a person *merely as a means* to maintain a tolerable condition. A human being, however, is not a thing and hence not something that can be used *merely* as a means, but must in all his actions always be regarded as an end in itself. (*G*, 4:429)

There are a few things to note about this argument, first of which is the context of the argument. It is important for the subsequent discussion to understand that Kant is arguing for a *perfect* duty not to commit suicide. As a perfect duty which constitutes a strict

obligation, there are no exceptions to the duty, nor are there genuine conflicts that could arise between this duty and other ends, even the ends adopted in the service of *imperfect* duties such as beneficence. Thus, in terms of stringency, to commit suicide is comparable to committing murder – both are absolutely prohibited. Second, Kant has not considered suicide *per se*. Rather, Kant limits his consideration to an agent who seeks to avoid an intolerable condition;<sup>23</sup> the adoption of this discretionary end is key to our understanding of the action as violating the Formula of Humanity as well as the difficulties this derivation presents. Third, one might question whether one is *using* oneself as a mere means to achieve this discretionary end. In other words, it is not immediately clear how the act of suicide uses our rational will to further the end of a tolerable condition except in the ordinary sense of setting an end and adopting the means to it; but this ordinary use of our rational capacities would not constitute a violation of the Formula of Humanity.

Paul Formosa (2014, 54) clarifies Kant's position by highlighting the agent's failure to "treat" humanity as an end of appropriate worth:

To treat humanity *in* ourselves as an end in itself means that we should regard our rational capacities as having an incomparably higher worth than that possessed by any of our merely discretionary or desired ends. But it would be irrational to damage or destroy something of incomparably higher worth or status for the sake of something with a lesser worth or status.

Thus, according to Formosa's interpretation, in committing suicide an agent is not *using* his rational capacities in the sense of impermissible *manipulation*, but rather treating his rational capacities as if it were of inferior value to the end of tolerable existence he has adopted. Said differently, the agent is treating his rational capacities as something of

---

<sup>23</sup> Kant parenthetically considers some casuistical cases, such as putting oneself in danger to avoid a greater or certain evil. However, none of these cases is problematic, and he says that a full consideration of such casuistical questions must be dealt with by "morals proper" (*G*, 4:429).

comparative value instead of something of incomparable worth. This interpretation accords with Kant's own brief explication – suicide violates the Formula of Humanity *not* because the agent destroys himself *qua* animal being, but rather because the agent fails to *regard* his rational nature (his humanity) as an end in itself.

In the *Groundwork*, then, there appears to be a clear derivation of a duty against suicide which does not appeal to any further principles or moral content. The act of suicide *for a particular reason* (*viz.*, to avoid a “trying condition”) is itself sufficient to show an absolute violation of the Formula of Humanity.

As in the derivations in *The Doctrine of Duty*, Kant's apparent rigorism immediately raises questions and provokes counter-intuitive cases. We have already considered above Kant's own views concerning suicide under extraordinary circumstances and for noble ends, as well as historical and contemporary examples that challenge the characterization of such difficult casuistical cases as “exceptions.” All of these examples serve to impugn Kant's stringent prohibition of suicide and to show that, contrary to such rigorism, suicide may be, under the right circumstances, virtuous or meritorious. Moreover, the attribution of “virtuous” to these cases does not simply imply that we intuitively judge the act of suicide to be permissible in these cases *despite* the Formula of Humanity; rather, these cases offer examples of suicide that are virtuous *in terms of the* Formula of Humanity. In other words, these cases serve as rebuttal to the claim that suicide always debases morality or humanity itself.<sup>24</sup>

---

<sup>24</sup> Though I do not defend it here, I find that the cases considered support the strong view that not only does suicide not always debase humanity; but further, that suicide when done for the right reasons and under the right circumstances may in fact elevate humanity. The mother who sacrifices herself for her child, or Cato who commits suicide in order to preserve the Roman Republic, disclose the sovereignty of reason over all heteronomous ends and inclinations.

We must ask whether the derivation of the duty against suicide in the *Groundwork* is susceptible to the same casuistical challenge. However, the purpose of considering casuistical difficulties is not simply to show that there are exceptions; but, rather, to show that the solution to difficult cases reveals a deeper problem with the standard cases – with the derivation of the duty in general. A canonical example used to reveal Kant's undue stringency is the case of a soldier who jumps on a grenade in an act of heroic self-sacrifice.<sup>25</sup> An initial assessment of such an act judges it to be in violation of the Formula of Humanity insofar as the agent regards his life as inferior to, or of lesser value than, the lives and aggregate happiness of the soldiers (or perhaps civilians) he saves. Though many find such a sacrifice (and even the calculation of value) to be intuitively noble (though not obligatory), on a strict Kantian view the act might appear impermissible in virtue of the self-debasement of one's own humanity, which is of incomparable worth, in relation to others.

Formosa proposes a solution to Kant's apparent rigorism, and argues that supererogatory acts of self-sacrifice are compatible with Kant's proscription against suicide so long as the act is done for a "moral end." In the case of a heroic soldier, Formosa contends that "[t]his is clearly not a case of suicide since the relevant intention is missing and, according to Kant, 'it is the intention to destroy oneself that constitutes suicide'" (Formosa, 2014, 55). The absent relevant intention is the intention to *kill oneself*. In the case relevant case of heroism, the intention adopted by the soldier is one of

---

<sup>25</sup> One may note that the soldier who jumps on a grenade does not, strictly speaking, commit suicide in that the soldier does not directly will his death. This is an important distinction that I address below; however, an initial assessment of the soldier's act as impermissible in terms of the Formula of Humanity remains reasonable insofar as the soldier puts his life in extremely high risk for the sake of the lives of others and in doing so makes a quantitative judgment of his value relative to others, and thus debases himself qua rational and moral being.

*saving others*. If this assumed intention is conjoined with the belief that one's individual worth is incomparable – that is, not of quantitative value, but of equivalent worth to not only another agent, but any number of other agents – then sacrificing oneself is permissible (though not required). This permissibility is granted because of the intentional disanalogy with killing oneself to avoid an intolerable condition: “In this case, unlike the case of sacrificing yourself or killing yourself for the sake of a merely discretionary end, you do not destroy something of absolute worth for the sake of something of lesser worth” (55). Thus, by focusing on the intentions of the agent, and specifically the regard adopted for his own humanity, we are able to show how Kant's apparent rigorism actually conforms to strongly held and plausible intuitions.

This analysis is fraught with difficulties which I can only address cursorily. (Again, our objective is to use these examples to reveal a deeper, general problem with the derivation of a duty against suicide.) By distinguishing between permissible and impermissible acts of self-killing by appealing to the intended ends or *effects* of the act, Formosa is wading into the many controversies related to the Doctrine of Double Effect.<sup>26</sup> Essentially, Formosa is arguing that the soldier (and others who commit heroic acts of self-sacrifice) does not *directly will* his death; indeed, we may safely assume that any soldier who does jump on a grenade hopes that the grenade is a dud (and if he does

---

<sup>26</sup> Formosa's position cannot (yet) be considered equivalent to the Doctrine of Double Effect (DDE), which sets multiple conditions on judging an otherwise evil act as permissible. Following Mangan, these conditions include: (1) that the action in itself from its very object be good or at least indifferent; (2) that the good effect and not the evil effect be intended; (3) that the good effect be not produced by means of the evil effect; and (4) that there be a proportionately grave reason for permitting the evil effect (1949, 43; quoted in McIntyre, 2014). Formosa appears to have conditions 1 and 2 in mind: the act of jumping on a grenade may be seen as an act of *saving others* and this end may be intended, and one's death or endangerment need not be intended but only foreseen. By appealing to a difference in intentions, Formosa shares in some of the challenges facing the DDE.

not, this is morally problematic).<sup>27</sup> One conspicuous difficulty is that even if the soldier does not directly will his death, he does directly will to put himself in extremely high risk of death. The fact that this action is done for the sake of others does not, on a strict Kantian view, exculpate the agent insofar as the action still debases his humanity; this is because the reason for the risk-taking is (or appears to be) grounded in a quantitative judgment of the value of humans (i.e., a higher number of lives has more value than a single life) rather than a moral judgment of the worth of humanity. Thus, acting for the sake of others does not avoid the problem of adopting an illicit regard for one's humanity.

Another proposed difficulty concerns the requirement that the soldier jumping on the grenade do so without the intention of killing himself, but only the foreseeable knowledge that the act will result in his death, since, following Kant, it "is the intention to destroy oneself that constitutes suicide." Shelley Kagan (cited in McIntyre, 2014) has argued that this description fails to both adequately capture the motivation of the soldier and the moral heroism that we attribute to the act. For many, the intentional self-sacrifice of the soldier – that is, intentionally giving up his life for the life of others – is not only permissible; it is precisely what makes the act morally honorable.<sup>28</sup> This argument is intended to show that the Doctrine of Double Effect does not explain what is permissible

---

<sup>27</sup> This distinction has a long history that goes beyond the scope of this project. (The distinction, and the original formulation of the DDE, is attributed to Aquinas [*Summa Theologica* II-II, Q. 64 Art. 7]; for an explication of the principle and a survey of related issues, see McIntyre [2014]; see also Engelhardt [2000, 277f, 320, 327f], for a discussion of the DDE in relation to abortion, suicide and end-of-life decisions from a traditional Christian perspective.) For our purposes, the question is limited what implications this distinction has for the derivation of duty proscribing suicide and other perfect, negative duties.

<sup>28</sup> This objection has some similarities with the problem of supererogation that is raised as an objection to Kantianism. This problem is, generally speaking, the accusation that Kant's moral theory cannot account for actions that go "above and beyond" obligation – that all actions are either forbidden, permissible, or obligatory. For the foundational article on this topic, see Urmson (1958); for further discussion see Guevara (1999); and for a defense of Kant's position see Hill (1971).

about the soldier who jumps on a grenade or the person who uses lethal force in self-defense. This criticism applies equally well to Formosa's proposed solution to Kant's rigorism: the truly challenging case is that of a soldier who sacrifices himself *with* the intention of killing himself, which many still find not only permissible but heroic. In this way, Kant's rigorism is not avoided by Formosa's proposed solution so long as we can find a case of a hero who does directly will his death.

Leaving these initial rebuttals, there is a further, more serious consequence of this proposed compatibilism that is of direct concern to our present critique. Kagan motivates his view that the voluntary jumping on a grenade involves an intentional self-killing by comparing the act to that of shoving another person onto the grenade to prevent the death of others. He suggests that in this case we would certainly view the intention of the agent to be that of *killing* the third party, and not merely an intention of *saving others*. In one sense, Kagan is right – this is how we would *typically* construe the intentions of the agent, and this does, I think, problematize the explanatory power of the Doctrine of Double Effect. However, what is at issue here, and what Kagan's example highlights for our purposes, is not what intentions we would typically ascribe to another agent, but rather what intentions an agent does or may hold, and what role these intentions play in the derivation of duty. In other words, what is typical in our ascription of intention entails nothing about what is necessary with regards to intentions, and thus entails nothing with regards to what actions are absolutely proscribed. The intention to kill another or to kill oneself is not intrinsic to the act of shoving another onto a grenade or jumping onto a grenade, respectively. However, according to Formosa it is the intentional ascription that



provides the character of the action<sup>29</sup> and permits the derivation of the relevant duty.

Problematically, the converse is also true: the *absence* of the relevant intention prevents the derivation of the relevant duty. And this is true for both suicide and, it would seem, the killing of another: if the person shoving another on the grenade did have the intention of saving the lives of others (and, we must hasten to add, had the consent of the person shoved) then the act would become permissible in terms of the Formula of Humanity.<sup>30</sup> Thus, the permissibility or impermissibility of the action depends on the character of the action, which in turn depends on the intentional ascription. But what grounds the ascription of an intention?

This dependency on intentional content presents a challenge which may be summarized as follows: If suicide uses humanity as a means, not in the sense of a tool or

---

<sup>29</sup> See Aquinas (*Summa Theologica* II, II, Q. 64, art. 7): "Now moral acts take their species according to what is intended, and not according to what is beside the intention since this is accidental as explained above (Q. 43, A. 3)."

<sup>30</sup> We must include here the consent of the agent whom is pushed as a condition for the permissibility of pushing insofar as such coercion would violate a strict duty of right, which I have not subjected to Hegel's critique. However, with regards to a possible ethical duty of omission, we may be able to omit this condition (Kant does not include wrongful killing in the *Doctrine of Virtue*). This condition is supposed to be necessary for the permissibility of the action insofar as physical coercion violates the rational agency of the coerced agent. Per my present discussion as well as my discussion below concerning deception, however, the absolute character of this condition is questionable given the following conditions: (1) if the coercing agent does not adopt intentional disregard for the coerced; (2) if coopting another's rational agency is not a sufficient condition for demonstrating the impermissibility of an action. Denying the possible permissibility of pushing another on a grenade puts the interlocutor in the difficult position of holding that killing oneself (in the case of suicide) is not necessarily wrong (so long as the right regard is adopted), but killing another is wrong *not* because of the intended or consequent death but because of the physical manipulation that constitutes disregard. As outrageous as this conclusion may seem (giving moral priority to manipulation over death) the reasoning is not very different from Herman's (1993b, 128ff) justification of self-defense. Lethal (and maybe any coercive) self-defense is justified not because one will otherwise be killed, but rather because the aggressor in killing me is "discounting my agency." Herman does not consider the right of self-defense from an agent who does not adopt such disregard (e.g., an agent who is starting a fire that I am confident will seriously harm or kill me); however, if she holds self-defense to be problematic and her solution to this problem necessary, then the implications seem clear and dire: one may not kill another to avoid him killing me, but only to prevent the discounting of my rational agency.

manipulation, but rather uses humanity in the sense of treating it as something of comparative value, then humanity is used as a means insofar as the agent regards his humanity as disposable.<sup>31</sup> This implies, further, that it is this *regard*, and not the surd or “bare” action itself that renders suicide impermissible. Moreover, on Formosa’s interpretation and my analysis this kind of use understood as a form of regard is inescapably *intentional*. In other words, in order for the suicide to be impermissible, the agent must consciously elevate another end above the end of humanity, and sacrifice his humanity *for the sake of* this other end. Critically, the converse – as revealed by an analysis of casuistical cases – is also true: if the agent does *not* regard his humanity as possessing mere comparative value, then his surd act of suicide may be judged permissible. However, there is nothing in the act itself – that is there is nothing about the surd act of self-killing – that requires an illicit regard of one’s own humanity.<sup>32</sup> One can commit suicide for any number of reasons. Indeed, Thomas Hill (1991), in his defense of a modified Kantian prohibition of suicide, enumerates (on my count) eight different motivational categories for suicide, only four of which, on his assessment, are morally

---

<sup>31</sup> I use this term in a somewhat technical sense as meaning “available for use.” What is meant is that the agent is regarded as something that may be used as a means for some other end that is taken to have greater value.

<sup>32</sup> One might object to this claim: some might hold that the act of suicide itself debases or expresses disrespect for humanity. I have not considered this view in any depth, first, because it accords with Kant’s view in the *Groundwork* which is considered below, and second, and even more importantly, I consider the claim is straightforwardly false. Perhaps a brief analogy will help motivate my dismissal: Suppose that an agent were in possession of something he considered to have great intrinsic value. Perhaps he has made a remarkable scientific breakthrough, and gained immense insight into the workings of the universe and physical laws of nature. The agent, moreover, values this knowledge and achievement intrinsically and more than any other valuable thing. Now suppose further that the agent knows that his discovery, knowledge, and achievement is going to be exploited and perverted for reasons contrary to his desires and objectives and then destroyed. It seems to me conceivable – even obvious – that the agent could destroy his own achievement, and even his acquired knowledge, in order to avoid such consequences without thereby reducing his regard for the achievement or knowledge as something of great, even utmost value.

objectionable.<sup>33</sup> Crucial to this analysis is the claim that the intentions ascribed to the action belong to the agent performing the action; that is, the intention ascribed to the act is the intention formed and ascribed by the agent's own will (it is the agent's *own regard* for his humanity that matters, and it is the agent's directly willing a particular end that determines this regard). Consequently, the intentional ascription necessary to the derivation of the relevant duty is a product of what Hegel's calls "heteronomy of choice" (*NL*, 76; see chapter 1, 13f).

In this way the problem facing the Kantian position surpasses possible false-positives or the proliferation of permissible cases of suicide – the problem is more fundamental. If the foregoing is the right interpretation of Kant – that is, if suicide consists in treating oneself as a means only if one, in committing suicide, regards his humanity as having comparative value – then Hegel's emptiness critique may be extended to this derivation of negative duties by arguing that, as in the case of natural

---

<sup>33</sup> The motivational kinds are: impulsive suicide, apathetic suicide, self-abasing suicide, hedonistic suicide, suicide when human life is no longer possible, suicide to end gross, immediate pain, suicide based on self-regarding moral beliefs, and suicide for altruistic reasons (which Hill does not address). Hill finds (and thinks we will find) the former four motivations to be objectionable and develops a modified Kantian principle to account for these intuitions (1991, 86-95). With the foregoing analysis, Kant's own view comes very close to Hill's modified Kantian view: what matters in determining the permissibility of an action is the attitude one adopts toward oneself qua rational being, and not the "external act" itself.

It may be helpful to consider not only the diversity of possible motivations, but the common actual motivations people adopt when considering and defending voluntary euthanasia. The Oregon Public Health Division 2014 Report cites the three most common reasons for euthanasia as being "loss of autonomy (91.4%), decreasing ability to participate in activities that made life enjoyable (86.7%), and loss of dignity (71.4%)." According one doctor and advocate for euthanasia in California, the most common reason given for euthanasia is not physical pain and suffering, but "existential suffering" (Will, 2015). Joe Magro, who suffers from ALS, gives the loss of dignity as his reason for euthanasia – or if the law still forbids medically assisted suicide, another form of suicide by his own hand: "I do not want to be a burden on anyone. I do not want to be dependent on my family to eat, drink, wash or simply go to the bathroom. I will live as long as I can but once it gets to a point where I cannot live life in dignity I will commit suicide" (Balzan, 2016). These are, of course, a limited sample; but they do provide a picture of not only what motivates actual people to commit suicide, but also, and more importantly, forms of motivation that do not entail an intentional self-debasement of one's humanity.

actions critiqued in chapter 2, the duty not to commit suicide relies on an unacceptable *subjectivism*. To review briefly the position in that chapter, I argued that Korsgaard's solution to the problem of natural actions entailed what Hegel considers to be an unacceptable subjectivism. The conducting of a natural action does not depend on a convention, and thus the universalization of the maxim of a natural action does not undermine the means of performing the action. However, Korsgaard contends that if we ascribe particular intentions to the agent, then natural actions with these intentions will not be universalizable. Korsgaard recognizes, however, that much depends on the intentional description of the case; I expand on this caveat, and argue that because intentions are and can be idiosyncratic, first, not every impermissible natural action can be ruled out; and, second, and more damning, this idiosyncrasy confronts Hegel's critical contention that the problem with idiosyncratic intentions is not simply that there can *be* idiosyncratic intentions (that, for example we may sometimes encounter strange intentions); rather, the deeper problem is that ascription of intentions is left entirely up to the subject, and the subject can ascribe any intention to his action which accords with the good he knows in the situation; that is, the good he recognize and adopt as his end (see chapter 2, §I.B as well as *PR*, §140d). This, first, opens the door to permitting any natural action whatsoever. Second, this critique implies that proscribed and permitted actions which accord with our intuitions are judged as such due to the subjectively determined intentions of the action.

The Kantian position faces the same objection in the present case of deriving a duty against suicide from the Formula of Humanity: a duty against suicide *per se* is not derivable from the Formula of Humanity. Rather, a duty against suicide is derivable only

if a specific self-debasing intention is adopted *by the agent himself*. This implies that the action of suicide, an act which is presented as violation of a perfect duty to oneself, and is thus on the Kantian view comparable to murder, may be judged permissible if an agent ascribes an intention to his act according to the good he perceives in the action – whatever that perceived good may be. Correlatively, we cannot judge this action to be impermissible so long as an appropriate intention is sincerely held by the agent. In simple terms, the agent gets to choose what his ends of his action are, including good ends, and thereby gets to choose what acts are permissible or impermissible. This “heteronomy of choice” constitutes, in Hegel’s view, an unacceptable and devastating subjectivism.

I have thus far limited my discussion of negative duties to oneself to the duty against suicide because, one, I think it represents a difficult case and, two, it represents a fundamental view in Kant that gives an account of our deontological tendencies (to the extent that some of us have them) regarding suicide, even if it fails to account for all of our intuitions. That being said, the critique presented can perhaps more easily be captured by considering some of Kant’s more notoriously rigorous positions – such as his view on “carnal self-defilement,” otherwise – and now often without embarrassment – known as masturbation. In the *Doctrine of Virtue* Kant initially adopts the same justificatory trope used to prohibit suicide and derives a duty against masturbation by arguing that it is contrary to the natural end of “sexual love,” which he posits as the preservation of the species (*MM*, 4:424). However, Kant then moves away from this form of derivation and adopts an argument that parallels those in the *Groundwork*. His reasoning for this shift is that “it is not so easy to produce a rational proof that unnatural, and even merely unpurposive, use of one’s sexual attribute is inadmissible as being a violation of duty to

oneself" (*MM*, 6:425). However, a "proof" that masturbation is impermissible – and even exceeds "murdering oneself" – is readily available, according to Kant, when we consider how an agent by masturbating "uses himself merely as a means to satisfy an animal impulse." Kant argues as follows:

[By masturbating] the human being surrenders his personality (throwing it away), since he uses himself merely as a means to satisfy an animal impulse.... But unnatural lust [unlike suicide], which is the complete abandonment of oneself to animal inclination, makes the human being not only an object of enjoyment but, still further, a thing that is contrary to nature, that is, a *loathsome* object, and so deprives him of all respect for himself. (*MM*, 6:425)

Kant's reference to "surrendering" and "throwing...away" one's personality cannot be taken literally; one still is, of course, in possession of one's humanity and personality *qua capacity* when one engages in this sexual act. Instead, I think Kant is best interpreted to mean that when a person engages in masturbation, he surrenders his personality *practically*, such that his personality is not operative in the action, and he is "taken over," as it were, by animal instinct. In this way, such surrender is an extreme and complete *disregard* for one's personality; and this disregard is what makes the act of masturbation wrong.

In response, I will only reiterate the arguments developed above and confidently suggest that one need not (though I am not arguing that one *cannot*) adopt such disregard for his personality by the mere act of masturbation. One may, for example, keep one's happiness in view and also judge that one's sex life is an integral part of that happiness. With this considered judgment, a product (it should be noted) of one's humanity, one may thereby conscientiously and deliberately engage in masturbation with the intention of, for

example, “exploring one’s own sexuality.”<sup>34</sup> In this way, one does not *surrender* his personality, but rather acts in accordance with his personality. (One may, we may concede, lose a full capacity for clear rational thought in the midst of the act; but this is true of procreative sex as well as heart-stopping roller-coasters.) As in the more serious case of suicide, what matters is the intention one adopts and the regard one has toward his humanity or personality. However, by making this the crux of our moral judgment of an action that may or could violate a perfect duty, the Kantian position leaves itself open to the idiosyncrasy of intentions; and the idiosyncrasy of intentions conjoined with the self-ascription of intentions entails for Hegel, as stated above, an unacceptable subjectivism.

As with the derivation of negative duties in the *Doctrine of Virtue*, we should not under-estimate or over-estimate the consequences of the foregoing critique. The cases I have considered have important consequences in and of themselves in that they show the inextricable subjectivism in the derivation of particular negative and perfect duties. These cases are also, as I have said, intended as a heuristic; the analysis presented can be extended to other duties. However, this does not imply that no duty can be derived from

---

<sup>34</sup>Such attitudes toward masturbation have, I will hazard, become unremarkable at least among certain demographic groups in the U.S. and Europe. Though previously unaware, I cannot say I am surprised that there is such a thing as International Masturbation Month, which was established after Surgeon General Joycelyn Elders, under Republican pressure, was forced to resign in part because of her suggestion that masturbation be taught to school children as a means of preventing the spread of HIV/AIDS (Jehl, 1994). (Interestingly, this is perhaps another way in which we may adopt masturbation as a means to a virtuous end without “surrendering” or disregarding our personality – though whether children may do this is a more complicated matter.) An example of a contemporary, popular feminist perspective on the importance of masturbation as a means of exploring one’s sexuality (presented with greater boldness and absence of embarrassment than I am willing to repeat here) can be found online at *FeministTimes.com*; the relevant claim for our discussion is presented candidly: “Giving young people the information they need to understand how to start exploring their sexuality solo will equip them with the self-knowledge and confidence to move on to healthy and safe sexual relationships as adults” (<http://www.feministtimes.com/wank-masturbation-women-pleasure/>). For psychological scholarship in support of this view, see Coleman (2003); for a more general exploration of the subject and the various results of sexual self-exploration, see Hogarth and Ingham (2009).

the Formula of Humanity using the strategy adopted in the *Groundwork*. Interestingly, I find the most compelling derivation provided by Kant to be of a proscription of an action that many in contemporary society may find unobjectionable – namely, periodic inebriation just for the sake of it. Perhaps more seriously and even more unavoidable is a proscription against permanently damaging one's rational capacities. I, at least, cannot identify intentions for such actions that would not involve a debasement of one's humanity to an object of value instead of an end in itself that is of worth and due our respect.

## 2. *False-promising and deception*

In the *Groundwork* Kant keeps to the examples used to illustrate the Formula of Universal Law, and thus his argument for a duty not to lie is limited to the case of false-promising. However, we can and others do extrapolate from this derivation a more general duty; however, this more general duty is not, I think, a perfect duty to be truthful, as in *The Doctrine of Virtue*, but rather a duty not to *deceive*. This limitation agrees with both the minor premise of the derivation as well as the category of duty being derived. As to the former, Kant relies on the claim that the relevant forms of dishonesty uses an agent as a mere means. As to the latter, unlike the proscription against lying in *The Doctrine of Virtue* which was a duty to oneself *qua* moral being, the duty derived in the *Groundwork* is a duty owed to others.

With these differences in mind, Kant's argument is again brief:

[H]e who has it in mind to make a false promise to others sees at once that he wants to make use of another human being *merely as a means*, without the other at the same time containing in himself the end. For, he whom I want to use for my purposes by such a promise cannot possibly agree to my way of behaving toward



him, and so himself contain the end of this action. (*G*, 4:429)

This argument requires some explication about why, exactly, the false promise uses the other agent as a *mere* means. Like in the case of suicide, the false promise does not immediately appear to constitute an illicit use of the other agent as a means. One might ask, “In what way is the agent’s humanity (his capacity for choice) being used or manipulated that is relevantly (i.e., morally) different from any number of other licit uses of an agent’s rational capacities?” (The canonical example of licit use of another’s rational capacities is that of using a bank teller to conduct one’s business. When we do this, we use the teller almost like a machine not unlike an ATM. The relevant difference for Kant and Kantians is that the teller is consenting to – choosing without coercion – to be used like this and for these reasons. Variations of this example, and the latter condition, will be important in subsequent discussion.) The difference lies not in the use of the agent’s will per se, but in the cooption of his will for a purpose or end that is not his own. That is, what makes false-promising a case of using the agent as a *mere* means as opposed simply to a means is that the agent lied to cannot “possibly agree to my way of behaving toward him and so...cannot contain the end” towards which his will is actually directed because this end is unknown to him. Thus, as Wood reconstructs the minor premise, false-promising “circumvents that person’s rational agency.” According to Korsgaard, the agent cannot *assent* to the actual end or *choose to participate in* the realization of this end (1996, 138-139).

Importantly, this analysis does not concern non-deceptive lies and general untruthfulness that remained problematic under the derivation in *The Doctrine of Duty*. If you present a false-hood to me that does not deceive, then it is not the case that I cannot

assent to the end or choose to participate in its realization. Even if you intend to deceive me, if I am not deceived then my rational agency has not been circumvented (though the intended deception will still violate a duty to oneself, as in *The Doctrine of Virtue*).<sup>35</sup> On the other hand, deception per se is not what makes the false promise wrong. That is, it is not in virtue of the fact I was told a false-hood and believed that false-hood that I was used as a mere means; rather, it is because that *through* this deception my rational agency was coopted and used for an end that is not my own that the false-promise is wrong. Again, unlike the derivation in *The Doctrine of Virtue*, this derivation does not reject as impermissible a deception that does not circumvent another's rational agency, such as a deception used as part of a joke (though, *pranks* would seem to be impermissible).

Second, as with all the negative duties discussed, this duty is also perfect, and thus the prohibition on false-promising and similar forms of deception is absolute. Indeed, Korsgaard considers the impermissibility of relevant forms of deception to be fully unavoidable, and this leads her to concede – in contrast to her defense of the Formula of Universal Law – that the infamously problematic case of the murderer at the door cannot be resolved using the Formula of Humanity (or the Formula of the Kingdom of Ends). This leads Korsgaard to propose a two-tiered modified Kantian theory where the Formula of Universal Law is used for “evil circumstances” and the Formula of Humanity and Formula of the Kingdom of Ends serve as the “ideal which governs our

---

<sup>35</sup> On this point Korsgaard writes: “suppose it is the case that if the other person knows what you are trying to do and has the power to stop you, then what you are trying to do cannot be what is really happening. If this is the case, the action is one that by its very nature is impossible for the other [he that is receiving the promise] to assent to. You cannot wrest from me what I freely give to you; and if I have the power to stop you from wresting something from me and do not do it, I am in a sense freely giving it to you” (1996, 139).

daily conduct" (1996, 151).<sup>36</sup>

Lastly, along with Korsgaard it is also important to emphasize Kant's point that the wrongness of the false-promise is understood in terms of the *possibility* of the agent to agree to the false-promiser's behavior. This criterion must be distinguished from other criteria, including that of an agent who merely "*does not* or *would not* assent to the transaction," as well as an agent who simply "does not happen to have the same end I do." As Korsgaard explains, if possible assent to an end is construed in terms of "what we are likely to be annoyed by" then the test will be subjective and the results contingent. Likewise, if the possibility of assent cannot be understood in terms of what ends I desire; we may both share the same desire, but you may still use me as a mere means if you fail to include me in choosing to realize this mutually desired end.

At this point I want to take a detour away from deceit. In amplifying the preceding point Korsgaard (very briefly) presents an example that I believe will clearly illustrate Hegel's critique. Korsgaard's example involves stealing a gift that is intended for you. In this circumstance the victim of theft and you share the same end – namely, that you receive the gift. According to Korsgaard and Kant, this mutual desire for an end is not sufficient reason for stealing the gift; though we may both be "motivated by the desire that you have" the gift, and "I may care about you too much or too little to be annoyed by the theft....this must still be a clear case of you using me as a mere means" (1996, 138). Korsgaard is using this example to illustrate the point that holding the same end as the person perpetrating a deception, or other impermissible act, is not a sufficient

---

<sup>36</sup> This solution turns on resolving the rigorism of the Formula of Universal Law which Korsgaard thinks can be accomplished, at least with regard to the murderer at the door example. See chapter 3 for my rejection of this solution.

condition for making the act permissible. What is at stake is not the end achieved, but the possibility of the agent participating in the end achieved. In the case of the stolen gift, though I want you to receive the gift, by stealing the gift you have circumvented my rational agency and made it impossible for me to participate in the realization of the end.

I think there are intuitive reasons for why such a theft is not clearly wrong, even if it would technically use an agent as a mere means. For example, the theft might save you or the giver from embarrassment; the theft might protect you from revealing something personal and private; the theft might serve to prevent the giver from indulging in vicious bragging or self-aggrandizement (by, for example, giving a gift that shows off his wealth). Like all of the ostensible duties considered so far, there are numerous and various counter-examples that impugn Kantian rigorism. If we set the proposed action in the right circumstances – which also often involves de-contextualizing and anonymizing the agents – then the action strikes us as wrong; however, if we contextualize the action differently and particularize the agents, we often get different intuitive results.

However, these counter-examples can be set aside in light of a more effective critique. The example of stealing a gift provides an opportunity to show how the wrongness of the act does not in fact depend on the proposed criterion, but rather on a moral assumption implicit in the example. In other words, what matters are the normative expectations and recognized claims of the agents in the scenario.

The first stop on this detour is showing that circumventing and even frustrating another's rational agency is under many ordinary circumstances not objectionable. This is because the other agent has no *reasonable expectation* that their choice be included. Suppose I want to change the landscape in a nominally regulated public setting, and other

than your protest, the action is benign (e.g., I want to put up a rope swing over a river). In such a case, we have conflicting desires and we have adopted conflicting ends, but I may, nonetheless, *disregard* your desires and your adopted ends without using you as mere means or violating any perfect duty. I can even (to satisfy one of Wood's requirements) *frustrate* your end of sitting peacefully and viewing an undisturbed natural setting.

Though one may (or may not) consider me inconsiderate, I have not violated a negative duty derived from the Formula of Humanity; and I have not done so because I have not violated any right or claim you have over the natural setting or its modification.

We can readily get the opposite result, however, if we posit a counter-factual normative specification. Now suppose that you and I live within a convention (and I think this may be official law or custom) of "community consensus" where we must agree to any change or disturbance to the surrounding, community-owned natural areas. As in the previous case, I have bypassed your rational agency and did not allow you to choose or participate in the realization of my adopted end. But, it is only in the latter case that the action is objectionable. I think this conclusion is obvious and even truistic; however, it is an intentionally simplified case designed to isolate the elements in virtue of which any relevant kind of action is wrong.

We can now extend this analysis to stealing: If we live under a convention of community property, and I take something from the landscape (e.g., sand dollars) then I have stolen; I have taken something that does not belong to me but belongs to all of us. I may also have frustrated your subjective end, and I may have done so without including your rational agency in the realization of my end. In this case I propose that my act of stealing is wrong because you can, in virtue of our shared convention, reasonably expect

that your assent be granted and that your choice be included in the realization of my adopted end. However, if we again make a simple change to the convention, such that, for example, a kind of Lockean *proviso* holds, then I can disregard your ends but only to a certain extent; specifically, I can take what I like, but only so long as I leave some for you *and* I make use of the items. In this case, my limited “taking” is not stealing and it is not wrong precisely because it is unreasonable for you to *expect* to be included per the norm of our shared convention.

Returning to the case of stealing an intended gift, however counter-intuitive it may seem to us, the same normative contingency holds. In this case the theft of the intended gift is wrong (or seems wrong) only because there is a preexisting assumption that the agent giving the gift retains a right to the gift up to and until the gift is given. If this assumption is removed, then it is not at all clear that the theft (or rather, “taking”) is wrong, *even though* the criterion of circumventing the individual’s rational agency would still be satisfied. We can easily imagine, for example, a convention whereby intended gifts are declassified as exclusive private property, and the intended recipient acquires normatively determined claims over the gift. Such claims might include how or when the gift is delivered, and importantly, such claims might be grounded in normatively recognized considerations, such as an interest in maintaining social propriety, avoiding embarrassment, or maintaining privacy. In this way, it is not clear that stealing an intended gift is wrong at all; and if it is wrong, it is not wrong in virtue of circumventing another’s rational agency; it is wrong in virtue of violating specific but contingent normative conventions and expectations.<sup>37</sup>

---

<sup>37</sup> Kant gives independent justifications and conditions for private property, so one might respond that such a convention may be contrary to right (6:250 and 255-256). However, this does not bear on the

One might respond that under such a convention one could not actually *steal* the intended gift. This is right: since the intended gift can no longer properly be understood as private property, it cannot be stolen. However, it can be *taken*, and it can be taken while circumventing another's rational agency. And it is this circumvention that was presented as the justification for what made this act of theft wrong. In other words, the objector is right in saying that "stealing" cannot be done if the context permits "stealing." But this is really just Hegel's point: *taking without assent from another* can be done if the convention permits it.

How does all of this apply to false-promising? Like the above case of theft, there are numerous and various intuitive examples that challenge the criterion given for determining that an instance of deception is impermissible. Chief among these is, of course, the infamous case of the murderer at the door. However, this may be seen as an outlier, a rare exception. There are other, more mundane examples that are equally problematic even if they do not impugn the Kantian position as acutely. These examples might include (as already suggested) deception to protect one's privacy, or deception to avoid subsequent gossip; but also, paternalistic deception, and especially paternalistic deception in service of helping a friend realize his genuine as opposed to impulsive ends; other examples include self-sacrificing or protective deceptions, or deceptions grounded in legitimate partiality.

I am going to set these casuistical cases aside and, as in the above discussion, focus exclusively on cases that show how the impermissibility of deception relies on illicit normative assumptions, and not the furnished Kantian criterion. I do this by

---

point being made here. Here the case was given as an example of a violation of an ethical duty to another, not as an example of respecting another's property right.

showing that this criterion can be met in comparable scenarios without being objectionable. Consider the following three cases:

1. *Endangered pretend-customer*: In this case a person believes he is being followed and is in danger. He enters a bank to avoid the perceived risk. The bank has a strict “for customers only” policy and the woman believes that, even if he were to tell the truth, there is a reasonable chance that he would be forced to leave and face the perceived threat. As a solution, the woman pretends to be interested in opening an account, deceiving the bank teller in the process.
2. *Consumer advocate*: In this case, there is no threat or danger. Rather, a woman works as a consumer advocate; part of her job is investigating the customer service and customer options of competing banks (he is not from a bank). As such, the woman goes to various banks and pretends to be a customer or potential client, asking about different services and products, and evaluates the customer service.
3. *Flirtatious friend*: In this scenario, a man is at a party or gathering with a friend and becomes innocently (i.e., he has no malicious intentions) interested in another person at the party. An opportunity arises for flirtation: the person of interest joins a conversation with the man and his friend. With the purpose of impressing the person of interest and provoking general conviviality, the man suggests to his friend that he recount a story involving the both of them since he tells the story best. There has not been an opportunity to consult with his friend beforehand and his friend is unaware of the man’s motivations or intended end.

While these three cases differ in various details, I propose that all of the cases involve unobjectionable deceptions. The first case, like the murderer at the door, involves the presence (or at least perceived presence) of evil; indeed, evil may be present in more than one way. The perceived stalker is an evil, so perhaps is the overly stringent rules about customers that are wrongly given priority over the safety of a non-customer. Interestingly, the person deceived in this case is not the same as the person presenting the situational evil – he is not a murderer or stalker. So if evil does explain the permissibility of the deception, then by allowing evil to be only contextually present instead of instantiated by he who is deceived we have already expanded the possible cases of permissible deception



considerably. Still, one might think that this is another special case that turns on the presence of evil, which leads to a consideration of the second case.

The second case includes no evident evil. Rather, it is a mundane case involving a professional activity that is conducted routinely in modern society. One could rebut that the consumer advocate has a duty and responsibility to fulfill her job, but this duty cannot supersede a perfect duty such as the duty not to lie. Also, one might object that the teller is aware that he might be surveyed by covert consumer advocates, and thus by taking the job he is tacitly assenting to the end of the advocate. However, this assumes knowledge of covert consumer advocates and such knowledge may not obtain – the teller may be naïve, or less plausibly, the consumer advocate may be the first of her kind. More importantly, knowledge of possible deception and a willingness to participate anyway do not entail tacit assent to an undisclosed end. We all know that some people lie, and we even expect to encounter some liars; we may even know that certain specific people lie, and we may nonetheless agree to talk with them, but not know when they are lying. None of this entails our assent or excuses the lie. Still, for some the institutional expectation may constitute special circumstances; note, however, that what counts as exceptional circumstances has again expanded considerably. Still, one may contend that under ordinary circumstances failing to include another's choice in the realization of an adopted end is impermissible.

This brings us to the flirtatious friend. In this case there is no evil present and there is no institutional expectation which could explain the exception – it is an ordinary social interaction involving friends. It must first be pointed out that this case also differs from the others in that it need not involve deception; but this is by design: it isolates the

Kantian criterion and reveals its ineffectiveness. Remember, deception is not in itself the property in virtue of which a false-promise is wrong; the false promise is wrong, along with other vicious deceptions, because the act coopts the agent's rational agency and denies him assent and choice in realizing your adopted end. These same criteria are met in the case of the flirtatious friend: he has adopted the end of flirting with a woman by means of making a good impression on her and conversationally facilitating an enjoyable time. The friend whom he uses cannot assent to or choose to help realize this adopted end. If this example is accepted, the boundaries of unobjectionable use of another as a mere means, understood in terms of the Kantian criterion, have extended to ordinary, everyday life.

To summarize: the Kantian criterion can be satisfied in numerous and various cases, including mundane cases, without being objectionable. Like the critique developed in the preceding section, this conclusion inverts or at least strikes a balance between what is exceptional and what is general; the false-promising scenario may just represent a subset of lying cases that are objectionable and may not exemplify the way in which deception is generally wrong. Unobjectionable cases where the Kantian criterion is met are just as standard as the impermissible cases like false-promising. Further, this critique implies that this sub-set of impermissible deception cases is impermissible for reasons other than the Kantian criterion: more specifically, in the case of false-promising the deception is wrong because it circumvents the agent's rational capacity in such a way that also violates posited and operative norms and normative expectations – the deception deprives the agent of his property over which he has a legitimate, recognized claim. This last conclusion further entails the most radical consequence of Hegel's critique: even

standard impermissible cases like false promising – the cases that exemplify the function and fullness of the Formula of Humanity – may become permissible given the appropriate alterations to the relevant normative specification and expectations. Though there may be a limit to this ethical malleability (and I have discussed some duties that can be derived from the Formula of Humanity), any immanent doctrine of duty will supervene on *Moralität* integrated with our normative convention – that is, with *Sittlichkeit*.

#### IV. CONCLUSION

The extension of Hegel's emptiness critique proceeded by demonstrating that while the Formula of Humanity is content-full, it is not sufficiently full to derive or determine particular duties, and is thus an insufficient ground for an immanent doctrine of duty. The emptiness objection cannot apply to the Formula of Humanity insofar as this narrow objection concerns the procedural formality of the Formula of Universal Law. However, Hegel's emptiness critique is broader, contending that Morality, as represented by Kant and the Categorical Comparative, cannot command beyond the abstract imperatives of "do the *right*" and "promote the *good*." In this way, Hegel concedes the fullness of the Formula of Humanity insofar as this consists in abstract respect for and moral consideration of rational subjects as ends. The objection is that Morality can go no further.

As has been noted above, this insufficiency of Morality is not simply a matter of there being exceptions to a generally applicable rule. The exceptions, rather, demonstrate that independent and contingent normative content (e.g., principles, ancillary arguments, intentional regard) is necessary for the derivation of an immanent doctrine of duties. This

deficiency of Morality, moreover, indicates the need for integrating particular, socially, culturally, and discursively contingent content in the form of *Sittlichkeit*.

To put this more concretely, I have selected exceptions with which many readers will agree: I anticipate that there will be significant agreement, for example, that euthanasia is acceptable if one's physical quality of life is sufficiently diminished. Likewise, I anticipate broad consensus about the permissibility of circumventing another's will (e.g., in cases of deceit) under conventionally ordinary conditions. In other words, the examples I have considered many will see as *reasonable* exceptions. What is reasonable, however, is a product of moral judgment. The broad consensus we (may) have, so long as it is not grounded in pure practical reason, presents the problem of normative variation based in what Kant calls a "disgusting hodge-podge of patchwork observations and half rationalized principles" (*G*, 4:409). The consequences of this empirical co-foundation are dire. If agreement over actual duty is not based in reason, then: (1) we have no rational grounds for prohibiting those who act unreasonably – that is, outside of or contrary to our current consensus; and (2) our current consensus may shift, conforming to different observations and rationalized principles which we may currently judge unreasonable. What is reasonable is subject to extreme (though perhaps not total) variation of both contemporaneous and non-contemporaneous judgment.

Some may propose that the criteria of reasonableness is empirical. It is important, then, to distinguish the additional normative content required from the uncontroversial need for morally relevant empirical or "anthropological" knowledge. A common trope in contemporary Kantian literature is that some knowledge about the empirical nature of humans, which is contingent, is morally relevant and necessary for the application of pure

moral principle. Thus, for example, the fact that we rely on certain forms of communication and cooperation makes deceit morally relevant; also, the fact that humans are physically vulnerable (i.e., can be injured, suffer from pain, etc.) makes threat a morally relevant act. Without this general and contingent knowledge about human nature, the application of a priori moral principles – or practical moral judgment – is impossible.

The cases I have considered here, however, do not rely on ignoring or opposing this necessary condition of practical judgment. The relevant moral properties are present in each of the cases and nothing in the arguments depends on the agents failing to acknowledge or recognize what is morally relevant. For example, the flirtatious friend is not unaware of the general normative proscription against deception or the reason behind it, nor need he be unaware of effective communication as a necessary means for humans to realize their various subjective ends. Rather, the agent is knowledgeable of what is required and permissible within his convention (in this case, the convention of friendship), and thus *aware* that a normatively relevant category of action (i.e., deception or circumventing a rational will) does not in this instance constitute disregard for humanity. Generally, then, what is at issue in these cases is whether what is known to be morally relevant becomes morally required, and the ground for this normative judgment. In the *Doctrine of Virtue* this ground was an intermediate premise that required independent moral judgment and ancillary justification; in the *Groundwork* this ground was disregard for humanity which is either subjectively or conventionally contingent.

As was noted in the chapter introduction, the extension of Hegel's critique is also important insofar as some argue that the Formula of Humanity (as well as the Formula of the Kingdom of Ends) augments our rational deliberative procedure, and thereby rectifies

the deficiencies of the Formula of Universal Law, including its formal emptiness. It is worth noting, then, how the extension of Hegel's emptiness critique to the Formula of Humanity is relevant to such views. Briefly, the answer is that in order for the Formula of Humanity (or the Formula of Kingdom of ends) to successfully augment the Formula of Universal Law, it must be sufficiently content-full to determine when an action (or maxim of an action) fails to treat a person as an end-in-itself. The Formula of Humanity is not independently sufficiently content-full to make such particular determinations.

An influential example of this view, forwarded by Barbara Herman (1993) holds that the Formula of Humanity provides the rationale for *why* a formally impermissible maxim (i.e., a maxim determined impermissible by the universalization test) is substantively impermissible. That is, the Formula of Humanity provides the wrong-making properties of a maxim, and we can interpret the formal impermissibility of a maxim in terms of this wrong-making property or value. Herman argues that this concerted use of the different formulations of Categorical Imperative may be useful for “resolving difficulties in judgment and maxim construction introduced by puzzle maxims”; more importantly, this concerted use fulfils the didactic intent of the Formula of Universal Law and provides deliberative guidance to the moral agent in “morally complex circumstances.”

The relevant value that can guide our deliberations is the value, or more properly *worth*, of humanity as an end-in-itself. Formally impermissible maxims (such as the false-promising maxim) which fail to set rational nature as the final end are substantively impermissible. Using the value expressed in the Formula of Humanity allows us to investigate and “determine, in our maxims, when our treatment of others is not

compatible with their status as rational agents” (1993, 236).

The relevance of Hegel’s emptiness objection extended to the Formula of Humanity is clear: the Formula of Humanity independently does not and cannot serve as the rationale which explains why an action is impermissible. The Formula of Humanity only provides the abstract reason why an action (or maxim of an action) is wrong: if the action fails to treat the rational agent as an end-in-itself. However, as we have seen, which action falls short of this abstract value and why an action falls short is not immediately derivable via the Formula of Humanity. Either an intermediate premise, with ancillary justification, is required for this determination; or the intermediate premise of disregard must be satisfied. With regard to the latter, whether the condition of disregard is satisfied is subjective (in the case of suicide) or a matter of coopting another’s will (in the case of deception; this latter disregard accords with Herman’s own example). Cooption or circumvention of another’s will, however, was itself shown to be an insufficient ground for determining the impermissibility of an action; that is, one may circumvent another’s will without showing genuine, or wrongful, disregard. In short: the Formula of Humanity is not a rationale for why particular actions (or maxims of actions) are wrong, and the supplementary principles necessary for showing if and when an action is actually wrong are problematic.

The final cases of deception and circumventing rational agency are helpful in amplifying this last point. I suggested that there are numerous intuitively unobjectionable cases of deception that meet the Kantian criteria of circumventing another’s rational will; I further argued many such cases are mundane and need not involve extreme circumstances. The aim with these cases was to show that the Kantian criterion of

circumventing or coopting another's will was not sufficient for determining impermissibility of the deceptive action – other conditions must be met.

Extrapolating from these cases, I will only here propose that an integrated approach to ethics, which incorporates particular and contingent normative content as an additional condition which must be satisfied, provides a tentative solution to these cases as well as other problematic casuistical cases: deception is permissible or impermissible in the respective cases, even when the Kantian criterion is met, because such forms of deception accord or conflict with conventional exceptions; our moral judgment is grounded in whether or not it is normatively expected that one be sincere under such conditions. One may coopt a friend's will under those circumstances where the normative convention of *friendship* permits it; one may likewise lie paternalistically within those relationships or contexts where such exceptions to generally held moral precepts hold; more controversially, one may coerce or kill when the good reason for the action is recognized as such and does not entail disregard for another *qua* rational end.

Left undeveloped, this remains a pithy and unsatisfactory solution. It does, however, accord with Hegel's own pithy solution to the problem of how to be virtuous: "What a man ought to do, or what duties he should fulfil in order to be virtuous, is in an ethical community not hard to say. He has to do nothing except what is presented, expressed and recognized in his established relations" (*PR*, §150). This approach faces its own difficulties which are beyond the scope of this dissertation; most fundamentally, the solution must give an account of how we determine which conventional norms and which exceptions we should accept. But this is the solution Hegel proposes; a solution beginning with the integration of *Moralität* and *Sittlichkeit*.





## Chapter 5

### The Formula of Humanity, Morality, and the Limits of Positive Duty

This chapter concerns the determination and application of positive duties from the Formula of Humanity. This topic presents a particular challenge for the extension of Hegel's critique. Each of the previous chapters concerned the limits of *Moralität* and, specifically, limitations concerning the determination or derivation of particular, concrete duties by the Formula of Universal Law or the Formula of Humanity, respectively. Turning to the determination of positive duties from the Formula of Humanity there is a critical difference: Kant does not attempt to derive *particular* or *concrete* positive duties from the Formula of Humanity. Rather, Kant limits his determination of positive duties to general duties to oneself and to others which are directly entailed by the Formula of Humanity: we are obligated to pursue our own self-perfection as both embodied and moral beings, and we are obligated to love others practically, which, Kant will further argue, entails beneficence, gratitude, and sympathy. This difficulty is compounded by the fact that these general duties are plausibly entailed by the Formula of Humanity. That is, if we accept the Formula of Humanity as a formula of morality – that is, if we accept the Formula of Humanity as a moral principle such that we are morally obligated to treat every rational being as an end-in-itself – then the broad and generic conclusion that we must show practical love towards other rational beings and their happiness follows readily; that we must pursue our own perfection is perhaps more problematic, but still plausible.

Thus, it may appear that Hegel's emptiness critique cannot be extended to the derivation of positive duties in *The Doctrine of Virtue*. In a limited sense, this is correct: unlike the arguments in chapter 4, I cannot show how particular duties are not derivable from the Formula of Humanity if no such derivation is attempted. In another sense, however, Hegel's critical objective remains relevant: the overarching objective of Hegel's critique has been to show that the Moral perspective represented by the Categorical Imperative is an insufficient ground for determining particular, concrete duties – that is, even if we are able to determine duties, the determination that can be made falls short of what is required of a moral theory. This broader objective remains true for the positive duties derived from the Formula of Humanity in three different senses.

First, in Section II, I again consider Hegel's subjectivism objection. As has been seen in previous chapters, Hegel's emptiness and subjectivism objections are closely related and the latter has been considered part of what I have termed Hegel's emptiness critique; specifically, the subjectivism objection has "filled the gap" when the emptiness objection is inapplicable (e.g., in the case of natural actions). This close relation is maintained here. In addition, Hegel's subjectivism objection has direct relevance for a particular positive duty derived from the Formula of Humanity; namely, moral purity. According to Kant, we have a duty to ourselves, *qua* moral beings, to develop our moral perfection, and this includes moral purity or acting from duty alone. In considering Hegel's objection to this duty, I elaborate on an objection developed by Allen Wood concerning the emptiness of the moral will. I argue that Hegel develops a critique of Kant's moral psychology that presents the following dilemma: if moral purity – acting

from duty alone – is maintained, then our particular duty is determined subjectively; if particular duties are determined by Morality, then moral purity is undermined. This argument rests on the premise that duty must be motivated empirically – or by what Kant calls *pathological* feelings. The clear Kantian rebuttal is that this contention is false, and that reason can be a determining ground of the will through the mediating organ of *respect for the moral law*. Accordingly, I expand on the original objection by adopting and explicating some of Hegel's arguments concerning the metaphysical and theoretical tensions entailed by moral purity found in the *Phenomenology of Spirit*.

In Section III I turn to the emptiness objection proper and present two arguments for the insufficiency of the Formula of Humanity to determine particular, concrete duties. First, I argue that as an entailment of chapter 4, which severely limited what constitutes our actual negative duties, the scope of what can count as a positive duty has expanded. Simply, if an action that was proscribed becomes permissible, then that permissible action can now be adopted as, for example, an act of beneficence. Alternatively, a formerly proscribed action which, for example, involves circumventing the rational will of another, may be permissibly adopted as a means to a dutiful end such as self-perfection. As in chapter 4, the fundamental concern is not merely that there may be exceptions to a generally applicable rule that prohibits the adoption of certain kinds of actions; the pursuit of positive duty; rather, the true concern is that no consistent and tenable principle has been provided for making such determinations.

The second argument in Section III also concerns the insufficiency of the Formula of Humanity for determining particular duties. This insufficiency has been shown in different ways corresponding to negative and positive duties: in the case of the Formula

of Universal Law and negative duties (chapters 1 and 2) illicit content was required to generate successfully a contradiction. In the case of positive duties and the Formula of Universal Law, the *absence* of particular content generated false-positives; that is, the universalization procedure proscribed permissible and obligatory maxims. In the case of negative duties and the Formula of Humanity, I argued that either independent *moral* premises are needed to derive particular duties or that the intermediate premises in conjunction with the Formula of Humanity are insufficient for deriving particular duties. Turning again to positive duties, the argument is comparable to the argument presented in chapter 3: namely, the Formula of Humanity does not sufficiently determine our positive duties, and therefore *allows too much*. Said differently, because the derivation of positive (imperfect) duties from the Formula of Humanity is indeterminate, the subject may choose from a manifold of possible valid duties what is his actually duty. This subjective determination of actual duty, according to Hegel, allows for the elevation of the “particular over the universal.” In other words, the subject may act for his own interests contrary to the actual right and the good – that is, he may act opprobriously – even though his actions are compatible with abstract Morality. This elevation of the particular over the universal, according to both Kant and Hegel, constitutes *evil*.

### I. POSITIVE DUTIES IN *THE DOCTRINE OF VIRTUE*

A detailed explication of positive duties derived and discussed in *The Doctrine of Duty* is unnecessary here since this task has been thoroughly undertaken and realized in the secondary literature already (see, e.g., Aune, 1979; Wood, 1999 and 2008; Guyer, 2006). Rather, the goal here is to provide a taxonomy of the duties derived by Kant and, drawing

from the secondary literature, to distinguish these duties from their more ordinary or familiar counterparts.

Unlike negative duties, there is no syllogistic derivation of positive duties from the Formula of Humanity. This is due, again, to the disparate levels of specificity: the positive duties derived are general and follow directly from a consideration the Formula of Humanity. Positive duties to ourselves and others start with an abstract duty of promotion which is directly derived from – or perhaps equivalent with – the Formula of Humanity. Because humanity is an objective end-in-itself we should promote or act for the sake of this end to the fullest extent. This latter addition – “to the fullest extent” – implies that we should not only promote our and other’s rational natures as we find them, but should act also so as to realize fully rational nature: “the promotion of the self-existent end means that we should let our rational nature be completely realized” (Lo, 1981, 192).

Gaining in specificity, Kant argues that this realization of humanity involves several duties which (like negative duties) correspond to ourselves as natural (animal) and moral beings: First we have a duty to cultivate our natural powers. To reiterate, our duty is not limited to the preservation of the “innate scope” our natural capacities, but includes the full realization of these powers, the scope of which is determined by our freedom. Kant reasons directly, contending that we “owe” these capacities to ourselves *qua* rational being. To emphasize, we do not pursue our natural capacities for their own sake, nor for the advantages that we may gain through their development, but rather as a “command of morally practical reason” in order “to be in a pragmatic respect a human being equal to the end of his existence” (*MM*, 6:445). In other words, I have a duty to

make my sensible self 'equal to the task' of my rational self, which is an end in itself, insofar as my sensible self is the vehicle or means of my rational and practical freedom. These natural powers include our "powers of spirit" and our "powers of soul"; the former being "those whose exercise is possible only through reason" including "logic, mathematics, and the metaphysics of nature," while the latter are those powers "at the disposal of the understanding" and include memory and imagination and other capacities that take "experience as their guide" (*MM*, 6:445). Though our duty to cultivate our natural powers has some specificity, no particular duty is determined (e.g., the duty "every rational agent has a duty to learn formal logic" is not derived) and the duty is left *wide*: "Which of these natural perfections should take *precedence*, and in what proportion one against the other it may be a human being's duty to himself to make these natural perfections his end, are matters left for him to choose in accordance with his rational reflection about what sort of life he would like to lead and whether he has the powers necessary for it" (*MM*, 6:445).

Second, also following from the duty to realize our rational nature to the fullest extent is the duty to perfect our moral selves. Perfection in this context has two senses: first, we are to perfect ourselves by acquiring a pure motivation such that the moral law is "by itself alone the incentive, even without the admixture of aims derived from sensibility" (*MM*, 6:446); that is, our actions must not only conform with duty but be done from duty. Again, the reasoning here is not syllogistic, but follows directly: if our humanity which is a practical and "inner freedom" is to be fully realized, then it must manifest itself in the world through the realization of actions without the corrupting influence of pathological incentives; correlatively, if pathological incentives are not to

corrupt our motivations, we must develop self-control or mastery over our inclinations. This rational mastering of one's self in turn involves the cultivation of moral virtue, the striving for moral perfection in one's "motivation and performance of all duties" (Lo, 1981, 192). In Kant's words, this realization means that "a human being has a duty to raise himself from the crude state of his nature, from his animality, more and more toward humanity, by which he alone is capable of setting himself ends" (*MM*, 6:387). Crucial to the discussion below, Kant does not hold that we have a duty to attain or instantiate moral perfection, but only to strive continually for and progress toward it: "It is a human being's duty to *strive* for this perfection, but not to *reach* it (in this life), and his compliance with this duty can, accordingly, continual progress" (*MM*, 6:447). The second sense of perfection pertains to doing our duty comprehensively; we must "be perfect" by fulfilling all of our duties and "attaining completely one's moral end with regard to oneself" (*MM*, 6:446).

With regard to positive *imperfect* duties to others, the argument again proceeds from the generic to the more specific. If humanity is an end-in-itself, and we are to act for the sake of humanity itself, then we should do so wherever we find humanity. Humanity is not found only in ourselves, but in all rational beings. It then follows that we should act for the sake of humanity found in myself as well as in others; or, we have a duty to act for the sake of others *qua* instantiators or humanity just as much as we have a duty to act for the sake of humanity found in ourselves. In still abstract terms, Kant says that this constitutes a duty of practical love; such practical love, Kant states, can "also be expressed as the duty to make other's *ends* my own (provided only that these are not



immoral)" (*MM*, 6: 450). In this context, the ends of others are not the same as the objective end of humanity, but rather the subjective ends adopted by each individual.

It is not immediately clear why the Formula of Humanity would establish a duty of practical love such that I am obligated to make the subjective ends of another rational agent my own; that I should *respect* another's subjective ends is clear – to frustrate, interfere with or undermine the subjective ends set and pursued by the rational will would constitute a direct restriction or neutralizing of another's practical freedom. However, why I am obligated to adopt these same ends does appear to follow from this duty of respect. Kant's argument is not much help. Instead of demonstrating how the duty to adopt another's subjective ends follows from the Formula of Humanity, Kant reprises and modifies his argument from the *Groundwork*, arguing that if we were to universalize *and make public* a maxim of indifference (i.e., a maxim to not act beneficently), this universal maxim would conflict with our willing to receive assistance when in need (*MM*, 6:453). Paul Guyer (2006), however, is helpful in relating this argument back to the Formula of Humanity:

one's interest in having the help of others available as a means to the realization of one's own particular ends is not a mere matter of desire for one's own happiness, but also a matter of respect for one's status as a human being who freely chooses ends and attempts to realize them in a rational way. (257)

Guyer seems to be arguing that respecting oneself as a practically free person entails an interest in the means that may be necessary to the activity of my practical freedom, which is the setting and realizing of subjective ends. Again, if I adopt my humanity as an end, I thereby adopt the realization of this end, which for embodied beings requires natural means, including the aid of others. Extrapolating from this analysis, if respect for oneself *qua* rational being entails an interest in the necessary means for realizing subjective ends,

including means in the form of aid from others, and I am obligated to respect all other rational beings as I respect my self *qua* rational being, then I must take in interest in – or adopt – aid for others, including my own practical love, as means to realizing fully their humanity.

Gaining in specificity, Kant identifies three additional duties related to practical love: primary among these more specific duties is *beneficence*, which is practical love applied to promoting “according to one’s means the happiness of others in need” (*MM*, 6:453). The second duty Kant identifies is *gratitude*: “Gratitude consists in *honoring* a person because of a benefit he has rendered us” (*MM*, 6:454)). Why or how gratitude is a duty, as opposed to a natural or expected response to receiving a benefit, is also not immediately clear. Kant emphasizes that this response is not prudential, but is rather “a direct constraint in accordance with the moral law” and indeed a sacred duty which “cannot be discharged completely” (*MM*, 6: 455). In other words, we ought never stop being grateful no matter what we do to show our gratitude. Kant’s argumentation is sparse: he contends that violating the duty of gratitude “can destroy the moral incentive to beneficence in its very principle” (*MM*, 6: 455). This argument remains obscure. It may be that Kant is contending that a universal absence of gratitude would erode the moral incentive in others to be beneficent, thereby doing away with beneficence which is our primary duty. In one sense this would be an odd position for Kant insofar as he insists that we have a duty to moral perfection (as already discussed) which obligates rational agents to act from duty and not because of returned gratitude. On the other hand, Kant does recognize our human frailty and that such moral purity is something to be striven for and not something which is always achieved; thus, realistically, gratitude may be a

condition for a consistently available moral incentive. In this way, our duty to gratitude would be indirect. This indirectness is shared with, and thus corroborated by, the third duty of love Kant identifies: *sympathy*.

Discussing sympathy affords the opportunity to emphasize the non-affective character of duties of love. Our duty of love – including all the secondary duties of beneficence, gratitude, and sympathy – is a practical duty; meaning that we have a duty to act and not a merely a duty to feel a certain way toward another. Thus, Kant contrasts beneficence with *benevolence* which is merely feeling satisfaction at the happiness of others, and does not consist in adopting the happiness of others in the fulfilling of their subjective ends. The duty of sympathy may seem to diverge from this practical characterization, but it does not. We do not, according to Kant, have a duty to feel sympathy *per se* – indeed, Kant disparages what he calls “unfree” and “communicable” sympathy, or compassion, which increases “the ills in the world” and cannot possibly be a duty *in itself* (*MM*, 6:457). Rather, we have an indirect duty to sympathy insofar as it serves as a means to realizing our primary duty of beneficence: “Nature has already implanted in human beings receptivity to these feelings. But to use this as a means to promoting active and rational benevolence is still a particular, though only a conditional, duty” (*MM*, 6:456). In this way both gratitude and sympathy are indirect, but also secondary to beneficence, which is our primary specific duty for realizing the general duty of practical love.

In addition to explicating our positive duties, Kant also emphasizes what duties we do not or cannot have. According to Kant we cannot have a duty to promote another rational agent's moral perfection insofar as moral perfection consists in a completeness

and motivational purity that an agent himself achieves (*MM*, 6:387). This is a questionable limitation; though one may not attain moral perfection for another, it does not follow from this that one may not aid in another's achievement, and thus have a duty to help cultivate the habits and dispositions that will enable another to achieve moral perfection. Kant likewise argues that we cannot have a duty to pursue our own happiness insofar as we are naturally disposed to pursue our own happiness, and conceptually there cannot be a *duty*, which we necessarily perform reluctantly, for what we are naturally disposed to do (*MM*, 6:386). Crucially, the duty we do not have regarding our own happiness is the *promotion* of our happiness; as will be discussed below, this does not rule out permitting or ensuring our own happiness.

Guyer observes that this restriction regarding the promoting of our own happiness is also implausible: though we may be naturally disposed toward our immediate happiness, rational agents are not always or consistently disposed toward their overall happiness – we often act in ways contrary to our general well-being. Thus, we may have a duty, for example, to forego immediate happiness for general well-being and we may pursue our general well-being with (at least concurrent) reluctance. Kant's position is also complicated by the fact that we do have a duty to avoid self-deprivation; that is, we are obligated to *not* deprive ourselves “of what is essential to the cheerful enjoyment of the pleasures of life” (6:452) nor are we to discipline our natural inclinations excessively. The former “deprivation” may take the form of *avarice*, or going beyond what is essential for enjoying the pleasures of life; the latter deprivation we might call *asceticism*. Thus, though we do not have a direct duty to pursue happiness, we are obligated to restrain any inclination that might excessively restrain our natural disposition to happiness.

As stated above the foregoing is not intended as complete explication or critical analysis of the positive duties Kant identified in *The Doctrine of Virtue*. Rather, it is a taxonomy and brief clarification of those duties which will help facilitate the extension of Hegel's critique. The first step in developing this extension is a discussion of Hegel's objection concerning the emptiness of the moral will and the possibility and implications of moral purity. This unique objection, interesting and important in its own right, is of particular relevance to our extension of Hegel's critique as it intersects with, and on Allen Wood's view undergirds, Hegel's subjectivism objection discussed in the preceding chapters and essential to the extension of the emptiness objection is Section III.

## II. THE EMPTINESS OF THE MORAL WILL, MORAL PURITY, AND SUBJECTIVISM

As we have seen in previous chapters, Hegel's emptiness critique is not limited to the standard emptiness objection but involves a constellation of arguments. One of those arguments concerns subjectivism, or the contention that subjective determination of intentional content allows for opprobrious maxims to be judged permissible in terms of both the Formula of Universal Law and the Formula of Humanity. Thus far, this objection has been considered independently – specifically, it has been appropriated from Hegel's larger critique as a critical response to Kantian rebuttals and solutions to perennial problems (see, esp., chapter 2). However, the problem of subjectivism can also be seen as a consequence of Kant's moral psychology and, more specifically, a particular Kantian requirement of moral motivation.

The relation between subjectivism and moral motivation is not new territory, and I will follow Allen Wood's presentation of the initial objection. However, as with Hegel's emptiness objection, there are elements of Hegel's case against Kant's moral psychology which are not incorporated into this objection; by revisiting Hegel's objection I aim to connect different arguments and, thereby, present more fully the difficulties facing the Kantian position. These further objections, moreover, specifically target the duty of moral purity.

### **A. Identifying the Case against Moral Purity**

In his book, *Hegel's Ethical Thought*, Allen Wood (1991) dedicates a chapter to the emptiness of the moral law; the early sections of this chapter address the emptiness objection specifically, where Wood first considers and rejects what I have called the "standard interpretation" (see chapter 1); Wood then develops a related objection by considering critically the derivation of the Formula of Universal Law and how Hegel's emptiness objection may be understood in terms of the failure of this derivation. In the last sections of the chapter Wood considers the "emptiness of the moral will" which he contends may be interpreted in terms of the failure to derive the Formula of Universal Law.<sup>1</sup> This interpretation of Hegel's objection, moreover, involves the explication of a dilemma facing Kantian moral psychology. In barest form, the two horns of the dilemma are (1) the acceptance of moral purity and (2) the incorporation of "sense-nature" or pathological inclinations and impulses into our moral psychology. The first horn leads to

---

<sup>1</sup> Wood writes: "If Hegel's emptiness charge may be interpreted as the denial that any contentful moral principles may be derived from the mere concept of a categorical imperative or objective ground, then it might also be interpreted as the denial that any contentful moral principles may be derived from the idea of acting from duty" (1999, 167).

the undermining of Morality in the form of unacceptable subjectivism; while the second horn leads to the undermining of Kantian duty and moral worth. Each horn – more importantly – moves the dialectic forward towards *Sittlichkeit*.

Before turning to this objection, I must provide a caveat. The objection that Morality is facing a dilemma and, specifically, an unacceptable subjectivism is drawn in part from Hegel's earlier works, including the *Phenomenology* and *Faith and Knowledge*. So also, is my augmentation. Moreover, some of the specific arguments raised against Morality, and thus Kant's moral theory, were originally levelled by Hegel against Fichte. As such, Hegel's specific arguments may not precisely parallel to his arguments in his later works, including the *Philosophy of Right* and the *Philosophy of Mind*. Even so, I believe Hegel's arguments from these earlier works may be appropriated and effectively levelled against Kantian Morality. Doing so, moreover, follows the influential explication of Hegel's objection by Allen Wood which has been taken up in the Kantian literature (see Allison, 1990; Ameriks, 2000). Nonetheless, where appropriate I provide corroborating textual support from the later *Philosophy of Mind* and *Philosophy of Right*.

Turning now to the objection: A crucial element of Kantian morality is that the agent does his duty for the sake of duty, or from respect for the moral law. As discussed above, it is one of our positive duties to strive for moral purity; that is, put conversely, to eliminate “more and more” all pathological motivations from the performance of our duty. Wood interprets Hegel as arguing that this moral purity is erroneous: whenever we perform our duty we are doing a particular duty, and the particular content of our duty is inextricable from our motivation. Hegel is not merely contending that moral purity is an ideal which we will never fully realize – which is Kant's own position; rather, Hegel

contends that being pathologically motivated (motivated by “empirical” inclinations and impulses) is a necessary aspect of performing one’s duty.

According to Wood, this contrary position may<sup>2</sup> be seen as based in Hegel’s more general view of action: “For Hegel, the relevant question of moral psychology is not about an act’s psychic causes but about the intentional descriptions that provide the agent with the (internal) reasons for doing it” (1999, 171). That is, the relevant question is different for Hegel because his view of duty and moral worth is not concerned with moral motivation *per se* – that is, Hegel does not determine that an action is done from duty by determining what causes an agent to act and indeed “denies that the will stands in any causal relation to its motives or to the other circumstances that determine it” (150).

Rather, Hegel’s theory of action, which informs his understanding of duty and moral worth, focuses on the intentional description which may be ascribed to the acting agent: to act “from duty” means only that duty is the predominate reason for why the agent acts (see Wood, 1999, 149). Thus, in Hegel’s theory of action, duty may be one’s reason for action, but that something is one’s duty *per se* is not the cause or ground of action.<sup>3</sup> This does not mean that reason for action and motivation are unrelated: adopting a particular end (including an obligatory or dutiful end) marshals an agent’s dispositions, inclinations, and habits – in a word, the agent’s character – to the achievement of that end.<sup>4</sup> Again, this

---

<sup>2</sup> Wood says that Hegel’s thesis is “ambiguous” (1999, 170). Different passages (some which I will consider below) support claims of different strengths.

<sup>3</sup> This position is an element of Hegel’s overall view on freedom and the will which cannot be considered in full here. For detailed accounts see Pinkard (2012, chap. 3), Pippin (2008, chap. 5); for an overview of Hegel’s position, see Bykova (2013).

<sup>4</sup> As Pinkard (2012, 98) explains: “For the concrete subject to act, she must be able as ‘substance’ to effect some kind of behavior, that is, exercise the causal powers of her animal nature. . . . Just as the intelligent animal will undertake its action on the basis of its own nature . . . the human agent will undertake its action on the basis of her own ‘nature,’ which is to be that of a ‘second nature,’ that is, an acquired set of dispositions, habits, and the like, and which will be constitutive of a set of basic orientations.”



is not a contingent fact for Hegel, but a necessary feature of action. On this point, Wood cites a particularly strong passage from Hegel's *Philosophy of Mind*, which I quote in greater length here:

An action is an aim of the subject, and it is his agency too which executes this aim: unless the subject were in this way even in the most disinterested action, i.e. unless he had an interest in it, there would be no action at all. . . . They [impulses and inclinations] are sometimes contrasted, on the whole to their disadvantage, with the morality of duty for duty's sake. But impulse and passion are the very life-blood of all action: they are needed if the agent is really to be in his aim and the execution thereof. The morality concerns the content of the aim, which as such is the universal, an inactive thing, that finds its actualizing in the agent; and finds it only when the aim is immanent in the agent, is his interest and – should it claim to engross his whole efficient subjectivity – his passion. (*PM*, §475)

In this passage we see a full integration of the agent's intention and empirical motivation. It is not simply the case that the agent is motivated by his empirical nature, though this is critical. In addition, the end, or here the "aim," of the agent does not become actual or the genuine aim of the agent until the inclinations and impulses of the agent are incorporated – the aim is actualized in the agent when the aim becomes "immanent in the agent"; that is, when the whole agent in the form of his "second nature" or character is oriented to the end. This is contrasted with the concern of Morality, which according to Hegel is the universal "content of the aim" and which is "inactive"; that is, a content – such as "this is my duty" – which remains abstract.<sup>5</sup> This view which incorporates the agent's own passions and interests in the determination of an end *qua* actual end is also seen in the *Phenomenology* where Hegel considers the reasons for action, including the performance of one's duty:

---

<sup>5</sup> A particularly strong statement of this view is given in a footnote (4: 401) in the *Groundwork* on the concept of respect (which I address in detail below). Kant writes concerning the end and interest of duty: "Any respect for a person is properly only respect for the law of which he gives us an example....All so-called moral *interest* consists simply in *respect* for the law."

I act morally when I am *conscious* of performing only pure duty and nothing else but that; this means, in fact, when I do *not* act. But when I really act, I am conscious of an 'other', of a reality which is already in existence, and of a reality I wish to produce; I have a *specific* purpose and fulfil a *specific* duty in which there is something else than the pure duty which alone should be intended. (*PH*, §637)

From both these passages we see, first, that on Hegel's view the causal ground of an action is found in the empirical nature of the agent; and second, that the agent sees himself – that is, he sees his *interests* – in the action performed and the end achieved and that the end is not actual until this interest is incorporated. This does not change in the context of duty and action done from duty: when an agent acts from duty he acts from rational necessity, but also sees his own interest (his own “satisfaction”) in the realization of this rational necessity. Thus, for Hegel, particularity, the specific features and elements of one's circumstance, are inextricable from actual duty – that is, the performance of duty in the world.

This perspective of moral action has severe implications for Kant's theory of moral motivation and moral purity. Wood quotes a passage from *Faith and Knowledge*<sup>6</sup> which presents the problem facing Kant:

The particular is something strictly empirical precisely because the emptiness of the pure will and the universal are what is truly *a priori*. So it would be contradictory to determine what right and duty are in and for themselves, for content at once suspends the pure will, duty for duty's sake, and turns the duty into something material. The emptiness of the pure sense of duty [respect] and the content continually get in one another's way. (*GW*, 183-184)

There is a contradiction because as soon as we determine duty in its particularity the duty *qua* end becomes empirical; that is, the duty as an end becomes something that we have an interest in and which we achieve through the marshalling of our character. But this is

---

<sup>6</sup> For supporting passages from the *Philosophy of Right* see §§121-124.

contrary to duty – the performance of my duty must be *from duty alone*, without any admixture of individual interest. Thus, the “sense of duty” – the pure ground for performing my duty – and determining the content of my actual duty inevitably conflict: if I determine my actual duty, I undermine the purity of my will and, thus, my duty; in determining my particular duty my duty is undone. Said differently: because the ends adopted by an agent will always be particular ends and my interest in *those ends* will always be inextricable from my reason for acting, if I determine that a particular action is my duty, then the particular, or empirical, features of that action will be included in my reason for acting, and thus be included as the grounds from which I act. However, if empirical reasons constitute the grounds from which I act, then I do not act from duty alone, but from particular ends, and thus abandon moral purity.

The contradiction involved in incorporating particular content *qua* duty – one horn of the dilemma – presents the Kantian with the problem of subjectivism – the other horn. If particular content (e.g., keeping a promise) is determined as part of one's duty, and the agent must act from duty in order to maintain moral worth, then the agent *must* act from the empirical elements included in his adopted end. Wood puts this point as follows:

Hegel thinks that in order to do one's duty as this particular duty, even if one derives the empirical features of that duty from some moral principle, one must act from those empirical features, and that to act from the empirical features of the act in this way is also to act from something that has the stamp of “particularity” on it; it is to act from empirical inclinations, interests, drives and passions. (Wood, 1999, 169)<sup>7</sup>

---

<sup>7</sup> Consider as an example the duty to help my mother as she ages. On Hegel's view, I may do this act *because* it is my duty. But my duty is particular, and on Hegel's view the presence of particular features does not keep it from being my duty, but actually those particular features *make it* my duty. I have a duty to help *my mother*. Performing this duty will involve particular empirical (or in Kant's terminology “pathological”) motivations; it will be part of my character to do what is rationally necessary, but also part of my character to help *my mother*. This kind of motivation is inexorable in terms of the cause of

As such, particular ends cannot be determined as one's duty – or, in order to maintain moral purity, duty must remain abstract. Of course, particular ends will be set – Kant and Hegel agree that ends are essential to any action. So, if Morality cannot determine particular ends as one's duty (e.g., if “helping one's mother as he ages” is not determined as a concrete duty), then the question arises regarding the determination of ends *at all* in the performance of one's duty. That is, Hegel contends that Morality only has the resources to consistently set abstract duty (do the right and promote welfare) and does not have the resources to consistently determine concrete duties. However, concrete ends in the performance of one's abstract duty must be adopted. So, the question arises concerning who or what sets these concrete duty-conforming ends.

According to Hegel, this non-determination of particular duty by Morality leaves it up to the subject to determine his concrete ends in the performance of his duty and to deliberate about which concrete ends are his actual duty. The following is from the same passage of *Faith and Knowledge*<sup>8</sup> cited above:

If we do not presuppose a genuine ethic, on the other hand, then since morality consists in the form, we are free to raise any moral contingency into the form of the concept, and thus to establish a justification and a good conscience for the unethical. As was shown earlier, the duties and laws in this system [of formal morality] make up an infinite, dispersed manifold every element of which is equally absolute. So the manifold makes a choice necessary. This choice [between duties] is strictly a subjective matter, for the objective aspect, the form of universality, is what is common to all [duties equally]. Now, we cannot think up any actual case of an acting that would not have several sides; for every intuition

---

my action. Or, again: my adoption of my duty to my mother *consists in* an interest to help my mother, as well as an interest in doing my duty, and provokes sensual motivations (love, loyalty, gratitude) in order to realize my duty.

<sup>8</sup> The subjectivity of determining particular duties by choosing some positive aspect of the action that “makes it a duty” is addressed at length in the *Philosophy of Right* (see §§133-140, and especially §140d). Moreover, I have discussed this subjectivism with regard to the Formula of Universal Law at length in chapter 2 (pgs. 57ff). Additionally, I discuss subjectivism and the emptiness objection as they pertain to the indeterminacy of positive duties below (pgs. 269-304)

of an actual case is infinitely determinable through the concept [i.e., abstract duty]. Some of these sides must constitute valid duties so that in obeying some duties, other duties will be violated and in violating some duties other duties will be obeyed. If the agent's own contingent bad sense determines this choice, then the sense is something unethical; but it justifies itself in its own sight and gives itself a good conscience through the awareness of that aspect of the action that makes it a duty. On the other hand, if the agent's disposition is decent enough to make him want to act objectively [i.e., to have a fully justified reason for action], he is confronted with the contingency of the duties [insofar as Morality has not and cannot furnish particular duties]. . . . So he is bound to fall into the state of sad indecision and weakness that arises from the following dilemma: there is nothing but contingency [in the options] open to the individual agent, and he cannot create any necessity by himself nor should he if he could. If he does decide for any one of the many duties, the possibility of the decision arises from his being unaware of the infinite mass of duties into which any actual situation for action can, and as a matter of duty must, be resolved; for such an actual situation, like anything that is actual, can be resolved into an infinity of qualities. The knowledge of these qualities which yield the duty concepts is impossible because the qualities themselves are empirically infinite. Yet, this knowledge is strictly required as a matter of duty. (GW, 184; underlining is mine)

We have different options for grounding duties objectively: a “genuine ethic” and Morality. In the absence of a genuine ethic we are left with Morality; however, for the reasons we have been considering, Morality cannot furnish particular ends as objective duties. Thus, it is left up to the subject himself to determine, from the manifold of possible ends that could be our duty, which concrete ends will, or will not, realize his abstract duty to do *right* and promote the *good*. If the subject has a “bad sense,” that is if the agent is disposed to those ends contrary to what is genuinely right and good, then the subject will do what is unethical but justify the unethical ends through an appeal to aspects of his action that give the pretense of duty (see also *PR*, §121 *Addition*). In other words, he will distort the description of the ends, and thus the action, to fit the good he sees in the action. The man of good disposition, on the other hand, is practically frustrated through indecision: he knows he must act objectively – he must do what is *actually* right and *actually* promotes the good. He also knows that he cannot determine

this himself – the actual good cannot be a product of his arbitrary choice. In other words, he recognizes that the various and conflicting ends can be differently and conflictingly described such that each may be construed as his particular duty; and further that this infinite variation of “quality” must be resolved before he can perform his duty with integrity – that is, without the false and subjective resolution which characterizes the man of bad sense. The knowledge furnished by this resolution is thus a necessary condition of duty, but not a condition that Morality can furnish.

We can already discern the direction this argument is heading: we must opt for the first alternative, a “genuine ethic.” In Hegel’s mature terminology, we must incorporate *Sittlichkeit* which provides the historical, cultural, and discursive particularity – that is, substantive normative grounds – necessary for determining our concrete actual duty. However, before we say that Hegel’s analysis leads us to a consideration of *Sittlichkeit*, must the Kantian accept the foregoing analysis? The dilemma facing the Kantian does not get generated unless we accept the claim that an agent must act from his empirical nature. If this is claim rejected, then Morality is free to determine particular duties without undermining moral purity.<sup>9</sup> Indeed, Kantians have supplied this rejoinder. Henry Allison (1990, 184-191) has responded to the “Hegelian critique” as presented by Wood and rejected the conclusion by rejecting the controversial, and thus far assumed premise, that an agent must act from the empirical grounds identified in the agent’s intention; that is, motive and end need not coincide. Allison writes: “On the contrary, it must be insisted once again that the centerpiece of Kant’s conception of rational agency is the Incorporation Thesis, that is, the claim that an incentive can determine the will only

---

<sup>9</sup> I am, of course, leaving aside here the emptiness objection, which argues independently that particular duties cannot be determined by Morality through a formal assessment, or test, of maxims.

insofar as it has been incorporated into a maxim" (1990, 189). Moreover, with regard to moral action or duty, Kant has an answer to Hegel's contention: namely, that the moral law may be an incentive to action by means of the recognition of the moral law and the consequent feeling of *respect*, which serves as a mediating organ that enables us to realize our duty and which does not have an empirical or pathological origin. Thus, a central premise has not been argued for, and the substantive Kantian position has not been addressed. As such, Hegel's view of action and, correlatively, moral worth has been presented merely as a competing view to Kant's with nothing to compel the Kantian.

In light of this deadlock of views, I seek to further Hegel's case by incorporating antecedent arguments from the *Phenomenology*. A full consideration of Hegel's arguments which precede and support the emptiness of the moral will objection cannot be considered here (as this would require a dissertation in its own right).<sup>10</sup> Rather, in advancing Hegel's case I focus on two components:

1. Metaphysical: To show that the Kantian position, even with an attenuated form of moral purity, requires a dubious metaphysical foundation and involves "metaphysical hypocrisy."
2. Theoretical: To show that the transition from moral purity to a duty to strive for moral purity generates theoretical tension with the effect of undermining the plausibility of the position as well as the normative force of the duty.

My reasons are as follows: The first component parallels well established and independent arguments in the Kantian secondary literature and connecting Hegel's arguments to these is worthwhile. The theoretical component of Hegel's arguments reveals an internal tension within Kant's theory that specifically targets the tenability of moral purity as a positive duty, and is thus of particular import for my present purpose:

---

<sup>10</sup> A survey and assessment of Hegel's objections in the *Phenomenology* as they pertain to Kantian Morality is given in Westphal's (1991) *Hegel's Critique of Kant's Moral World View*.

showing the insufficiency of Morality for determining particular duties. As noted in the introduction to this section, developing arguments from Hegel's earlier works in conjunction with arguments from the *Philosophy of Right* is disputable insofar as Hegel's early and mature views may not precisely coincide. Nonetheless, particular analyses from the *Phenomenology* may be appropriated and relevantly applied to the above argument concerning the emptiness of the moral will.

### **B. Hegel's Case against Moral Purity<sup>11</sup> – the Dissemblance of Morality**

In the *Phenomenology* Hegel gives an extensive analysis (§§622-631) of what he calls the “dissemblance” or “displacement” of the Moral perspective.<sup>12</sup> Hegel accuses the Moral perspective – or “moral self-consciousness” – of not being in earnest about key positions; of forwarding a strong commitment and then attenuating or abandoning that position by adopting conflicting positions. In this analysis Hegel, does not address Kant exclusively. Nevertheless, much of Hegel's critical analysis clearly has Kant in view and remains relevant for Kant's position concerning positive duties; specifically, the positive duty to progress toward moral perfection.

#### *1. Metaphysical hypocrisy*

---

<sup>11</sup> Throughout this section from the *Phenomenology* Hegel refers to “moral perfection”; however, he seems to have in mind what Kant refers to as “moral purity.” I maintain Kant's terminology in my own explication and argumentation.

<sup>12</sup> The title of this section already demands some commentary. As Beiser notes, the German *Die Verstellung* is derived from *verstellen* which can mean both “displacement” and “to disguise or conceal” – hence the translation as “dissemblance.” Beiser argues that Hegel is knowingly adopting this ambiguous language, but also that he gives his own sense of what *Die Verstellung* involves in this context: To avoid contradiction, Moral self-consciousness “shifts back and forth, first claiming that one term is essential and then its opposite. In doing so, it deceives itself, and when it really has not” (Beiser, 2009, 217).



Hegel begins his analysis with the primary dissemblance concerning the purity of the moral will:

Moral self-consciousness asserts that its purpose is pure, is independent of inclinations and impulses, which implies that it has eliminated within itself sensuous purposes. But this alleged elimination of the element of sense it dissembles again. It acts, brings its purpose into actual existence, and the self-conscious sense-nature which is supposed to be eliminated is precisely this middle term or mediating element between pure consciousness and actual existence – it is the instrument or organ of the former for its realization, and what is called impulse, inclination. Moral self-consciousness is not, therefore, in earnest with the elimination of inclinations and impulses, for it is just these that are the *self-realizing self-consciousness*. (PH, §622)

On the one hand, we may read Hegel here as again insisting on the incorporation of our empirical nature – our “sense nature” in the form of impulse and inclination – in the actualizing of our ends and simply identifying our sense nature as the mediating organ that realizes moral self-consciousness. If this is so, however, we cannot at this point agree with Hegel concerning the dissemblance attributed to Morality: the argument would seem to be that because the agent, in possession of Moral self-consciousness, *acts* and brings his “purpose into actual existence,” and this realization requires the mediating function of his sense-nature, then Moral self-consciousness has thus shifted from its true position. However, this position, as it stands, merely dismisses the Moral, or at least Kantian, position, which is that such realization of duty is accomplished with a pure will through the mediation of *respect*.

On the other hand, by identifying inconsistencies within Kant's theory of moral motivation we can read Hegel as forwarding a substantive objection. That moral-consciousness is not in earnest about moral purity is variously evident in Kant's theory of moral motivation. Kenneth Westphal (1991, 149-151) has argued that both Kant's appeal to the highest good as a “check [on] the influence of inclinations” and to sympathy as a

cooperating feeling in the realization of duty lead to the same conclusion of the insufficiency of duty alone as a motivating ground. Focusing on the latter appeal, we have already seen that Kant gives to sympathetic feelings a significant role in moral motivation, identifying them as an effective means to our moral end of beneficence. Westphal notes, in support of Hegel's contention, that this identification includes the recognition that sympathy may play an essential role. Kant writes of sympathy: For this is still one of the impulses that nature has implanted in us to do what the representation of duty alone might not accomplish" (*MM*, 6:457). Two initial observations about this striking passage: first, Kant clearly identifies sympathy as *one* of the impulses which nature has implanted in us, suggesting that other impulses may be for the same purpose of realizing our duty; second, Kant clearly identifies sympathy (and by extension, other impulses) as supplementing respect for the moral law, which as we shall see below, is a product of the representation of the moral law.

Westphal takes pains to distinguish the cooperative function of sympathy from acting from sympathy or another inclination, and in doing so deflects a Kantian rejoinder forwarded by Allison that sympathy merely "counteracts" other inclinations allowing duty to become the sufficient incentive to action. Westphal observes that while this model of moral motivation allows for duty to be "decisive" in our action, it does not make duty or the moral law the "sufficient determining ground of the will": "Determining ourselves to do something in these circumstances still requires the cooperation of an inclination, sympathy, in order that the balance of interest can be shifted by our allegedly pure interest in duty" (150).

Such cooperation is only possible if sympathy, and other impulses, conform with

duty, which Hegel contends Morality requires in a continuation of the same passage.

But also they [impulses and incentives] ought not to be *suppressed*, but only to be in *conformity with Reason*. . . . But impulse is not in fact merely this empty shape which could have within it a spring of action other than the one it is, and be impelled by it. For sense-nature is one which contains within itself its own laws and springs of action; consequently, morality cannot therefore be in earnest about being itself the mainspring of impulses, the angle of inclination for inclinations. (*PH*, §622)

To illustrate Hegel's point, consider, for example, the duty to not avoid places of poverty and illness in order to cultivate sympathy. If this is to be effective, my emotional character must respond to such sites in the appropriate manner such that sympathy and not, say, disgust or fear is engendered. This is to say that my impulses, if they are to be a means to the ends of duty, must develop in a manner which conform to the demands of Morality. This conformity, however, is a matter of empirical causation and luck. Finally, if this is the case, Kant's dissemblance has profound implications for his general theoretical commitments, including his commitment to absolute moral responsibility as grounded in autonomy, is undone.<sup>13</sup>

Hegel's objection has further implications for Kant's theory of moral motivation in light of his insistence on the humiliation of incentives by reason in the form of respect for the moral law. This is particularly the case if one reads this passage more broadly, as

---

<sup>13</sup> Thus, for example, Kant's position in these passages from the *Metaphysics of Morals* regarding the necessary cooperation of our empirical nature would seem, in light of Hegel's critique, to directly oppose his earlier and stringent view of moral responsibility expressed in the first *Critique*:

Now even if one believes the action to be determined by these causes, one nonetheless blames the agent, and not on account of his unhappy natural temper, not on account of the circumstances influencing him, not even on account of the life he has led previously; for one presupposes that the series of conditions that transpired might not have been, but rather that this deed could be regarded as entirely unconditioned in regard to the previous state, as though with that act the agent had started a series of consequences entirely from himself. This blame is grounded on the law of reason, which regards reason as a cause that, regardless of all the empirical conditions just named, could have and ought to have determined the conduct of the person to be other than it is. (*CPR*, A555/B583)

does Beiser, as confronting the contradictory causal relation between the moral will and “all motives of sensibility” (Beiser, 2009, 218) – that is, our *noumenal* and *phenomenal* natures.

Such a broader understanding not only makes relevant Hegel's objection vis-à-vis Kant's theory of moral motivation; it also brings it into conversation with other objections and analyses within the Kantian secondary literature concerning, what may be termed, Kant's metaphysical hypocrisy.<sup>14</sup> Various commentators (Wolff, 1973; Hinman, 1983; Henson, 1979) have noted the fundamental inconsistency and unintelligibility of Kant's strict position on moral motivation. Above I have noted that Kant and Kantians may appeal to the notion and affective<sup>15</sup> motivational organ of rational *respect* in response to Hegel and his insistence on the incorporation of impulse and inclination in the performance of one's duty; the positive feeling of respect, along with the negative effect the representation of the moral law has on our inclinations, allows for the moral law to be adopted as our incentive to action (see Allison, 1990, 120-126). Because Kantians may appeal to this model of moral motivation, I noted that Hegel's critique of Kant amounts to no more than a contrary account of moral motivation. However, the requirement of respect *qua* ground of moral motivation is problematic in light of Kant's metaphysical commitments.

Kant tells us that respect is the result of a two-stage motivational operation.<sup>16</sup>

---

<sup>14</sup> In referring to Kant's position as hypocritical I follow Hegel in another of his characterizations of *Die Verstellung* (see *PH*, §631).

<sup>15</sup> The exact nature of respect *qua* mediating organ of motivation is controversial. Some commentators defend an *intellectualist* interpretation of respect where the rational emotion of respect is not directly motivating. See, for example, Reath (2008).

<sup>16</sup> This two-stage operation seems to be misunderstood by Hinman (1983), who writes that the feeling of respect humiliates one's inclinations: “Thus ‘respect for the moral law,’ Kant writes, ‘must be regarded also as a positive but indirect effect of the law on feeling, insofar as this feeling [of respect] weakens

First, the moral law itself *humiliates* [*demütigt*] our inclinations:

All the inclinations together constitute regard for oneself (*solipsismus*). This is either the self-regard of *love for oneself*...or that of *satisfaction with oneself*. The former is called ... *self-love*; the latter *self-conceit*. Pure practical reason merely *infringes upon* self-love.... [b]ut it *strikes down* self-conceit altogether.... (CPrR, 5:73)

This effect on our inclinations, according to Kant, causes “what might be called pain.” It is critical to follow a description of this relation of humiliation with Kant’s description of *how* such humiliation could take place – especially as Kant (somewhat misleadingly) uses language such as *infringing on* and *striking down* psychological states that are phenomenal – that is, belong to the causal order – with reference to pure practical reason. As such, it must be emphasized that the humiliation does *not* occur through direct *causation*. Rather, Kant explains that the humiliation of our inclinations is brought about, or caused, by the *representation of the law*:

the representation of the moral law deprives self-love of its influence and self-conceit of its illusion, and thereby the hindrance to pure practical reason is lessened and the representation of the superiority of its objective law to the impulses of sensibility is produced and hence ... the relative weightiness of the law in the judgement of reason. (CPrR, 5:75-76)

In this way, our recognizing of the moral law as the supreme law makes way for the practical effect of the moral law (see Reath, 2008, 11).

Even so, this negative effect on our inclinations is not enough to show that the moral law is the sole and sufficient determination of duty and requires a second stage.

---

the hindering influence of the inclinations through humiliating self-conceit’ (255). However, if we look at Gregor’s 1996 translation we see the same passage identifies the law itself as humiliating self-conceit: “...insofar as the law weakens the hindering influence of the inclinations by humiliating self-conceit....” (5:79). Furthermore, this interpretation is corroborated by Kant’s earlier explication considered in the text (that is, CPrR, 5:73). Respect is a positive feeling that does serve as an incentive to obey the moral law that is generated in response to the humiliation of our inclinations by the moral law.

The “thwarting” of our inclinations and corresponding pain also has a positive effect:

But since this law is still something in itself positive...it is at the same time an object of *respect* inasmuch as ... it *weakens* self-conceit; and inasmuch as it even *strikes down* self-conceit, that is, humiliates [*demütigt*] it, it is an object of the greatest *respect* and so too the ground of a positive feeling that is not of empirical origin and is cognized a priori. (*ibid.*)

Respect, then, is the *affective* effect of the negative humiliation of our inclinations brought about by the moral law itself. This positive feeling, moreover, functions as a rationally grounded incentive to “compliance with the moral law” (*CPrR*, 5:79). Kant emphasizes, and we must follow him in repeating this point, that respect is not an *incentive to* the moral law; that is, it is not an incentive that exists prior to the moral law; but “instead it is morality itself subjectively considered as an incentive inasmuch as pure practical reason” through humiliating our inclinations “supplies authority to the law, which now alone has influence” (*CPrR*, 5:76). By this Kant means that respect is the subjective and sensible experience of the moral law itself.

On the one hand, the account of practical effectiveness of pure practical reason is a necessary consequence of Kant’s prior and fundamental metaphysical views; on the other, it presents a difficult problem. Kant famously distinguished between the *noumenal* and *phenomenal*, or alternatively the intelligible and sensible. For our purposes, there are two crucial facts related to this division: first, the phenomenal is governed exclusively by causal laws whereas the noumenal is not determined or subject to causal laws. Second, we as agents are not entirely noumenal or phenomenal: we are, after all, embodied beings that also possess the transcendental ego including the categories of the understanding, which, in part, constitutes the necessary conditions for any experience whatsoever – that is phenomenal reality. An aspect of us – our noumenal aspect – must necessarily be prior

to and independent of the phenomenal.<sup>17</sup> Moreover, as free beings we possess a noumenal will (*Wille*) as well as a noumenal faculty of “arbitrary choice” (*Willkür*); as such, every free act, even actions done when inclinations are the incentive, are a product of a noumenal act (see Allison, 1990, chap. 7.) Also important is that all of our emotions *save respect* are phenomenal and thus products of causal relations governed by causal laws. Given this metaphysical division and categorization of emotions (save respect), Kant has no option but to find a feeling and incentive “that is produced by an intellectual ground” (*CPrR*, 5:73).

Some authors see an immediate problem with this theory of motivation in light of Kant's metaphysical commitments. As Hinman observes,

The root difficulty which gives rise to such tension at this point in his analysis is that he has so framed the question of the efficacy of the moral law that he is required to provide an account which has both causal and non-causal elements. On the one hand, Kant must produce something which has causal effects sufficient to outweigh our natural inclinations; on the other hand, he must show that this same thing is not itself the effect of some natural cause. (1983, 255)

We might read this passage as arguing that Kant is committed to the moral law itself being both outside of the causal order but simultaneously having causal effects; that the “something” which Kant must produce which has causal effects and is not the effect of a causal chain is the moral law. This would clearly be contradictory. However, this

---

<sup>17</sup> This is perhaps an oversimplification. Paul Wolff (1973), in attempting to parse our various “selves” and to determine to which world they belong (*noumenal* or *phenomenal*) actually identifies four selves: the empirical, the *noumenal*, the transcendental ego, and the moral self. However, this complication does not undermine the central point that we have selves (or, as I would prefer, we are both phenomenal and non-phenomenal) fully independent of the phenomenal, and on which the phenomenal is conditioned. Wolff writes: “The entire world order, including both the objects of nature *and the empirical self that perceives them*, is produced by the mental activity of synthesis [of the transcendental ego]. The active, knowing mind or transcendental ego, knows things as they appear rather than as they are in themselves, including of course itself, which it knows only as appearance.”

objection cannot be right and must be put aside: as we have just seen, Kant does not give such a wholly contradictory role to the moral law itself. Rather, it is our representation of the law that operates within the causal order.

It may appear that Hegel's objection follows this same line, since he writes that "sense-nature is one which contains within itself its own laws and springs of action; consequently, morality cannot therefore be in earnest about being itself the mainspring of impulses, the angle of inclination for inclinations." That is, he appears to be arguing that morality (i.e., the moral law) cannot itself be the mainspring or causal "angle" of inclinations. However, I believe a more nuanced reading is possible. This reading focuses on the independence of inclinations, and our sensible nature generally, instead of the causal role of the moral law. Specifically, that the inclinations must be independently *receptive* to a representation of the moral law. Reath makes this point as well: "The very fact that the moral law can check the inclinations and 'humiliate' the pretensions of our sensible nature reveals our responsiveness to rational principles" (Reath, 2008, 9).

This receptivity poses a problem. Kant rightly observes that though sensibility is a condition for feeling respect, there can be "no *antecedent* feeling in the subject...attuned to morality" (*MM*, 5:75); that is to say, we cannot feel respect because we have desire to conform to the moral law and still say that respect is rationally grounded. If we were to feel respect because of an antecedent desire to conform to the moral law, then our empirical nature would be the cause and mutual ground of our moral motivation; our empirical nature would, in effect, be an antecedent condition of genuine duty. This problem need not be limited to *desire*: our moral purity, and thus genuine duty, is undermined if our empirical and causal nature constitutes an antecedent condition for



acting from duty. With this broadened problem in mind, we may reconsider Kant's acknowledgement concerning our responsiveness to the moral law.

This responsiveness is indisputably a product of a phenomenal, and thus causal, state of affairs. In other words, though we need not have an antecedent desire to respect or conform to the moral law, we must nonetheless be antecedently attuned to respond to or receive a representation of the moral law. Wolff gives a succinct statement of this objection:

It is a fact about my *phenomenal* character that consciousness of submission to a self-made law produces a feeling of reverence in me. As far as moral theory is concerned, this fact is no different from the fact that a malicious pleasure is produced in me by the awareness of my enemy's misery of the fact that a feeling of well-being is produced by the satisfaction of bodily needs. (Wolff, 1973)

The nature of the ground does not change the nature of the effect. My responsiveness in each of these cases (respect, pleasure, feeling of well-being) is a product of my phenomenal nature regardless of what I am responsive to. Returning to the conditions for respect, we must be *caused* to be responsive to the representation of the moral law and even specific instances of a representation of the moral law. If receptivity is a condition for humiliation by the representation of the moral law, then I am necessarily caused to be susceptible to such humiliation. This entails that, every time I act from duty, my empirical nature must antecedently cooperate. My causally contingent empirical nature, then, is a mutual ground of moral purity and duty. Conversely, if I am not caused to be receptive to the moral law, I could not act from respect for the moral law.

Indeed, Kant seems to recognize this possibility, at least conceptually, when he discusses moral feeling in the *Introduction to the doctrine of virtue*. Kant again distinguishes moral feeling from pathological feeling by attributing to each a different

origin: pathological feelings precede the representation of the moral law, while moral feelings succeed it. Kant also emphasizes that this feeling involves receptivity or, as he calls it in this passage, susceptibility. Kant further emphasizes that we can have no duty to be susceptible to moral feeling, and that the absence of susceptibility would entail that one is “morally dead”: “No human being is entirely without moral feeling, for were he completely lacking in receptivity to it he would be morally dead” (*MM*, 6:400). What Kant does not appreciate is, first, the requirement that such receptivity be antecedently determined and, second, the condition of constancy with regard to such receptivity and, hence, moral responsibility. In other words, Kant seems to disregard the empirically conditioned possibility of token instantiations of moral death.

Hegel’s objection may be qualified to accord with this form of the hypocrisy objection: we cannot say moral self-consciousness disassembles insofar as it, contradictorily, determines morality to be the “mainspring” of the inclinations if this means that the moral law directly causes the humiliation of the inclinations. However, in light of Hegel’s preceding statement, that inclinations, and our broader sensible nature, is not “this empty shape which could have within it a spring of action other than the one it is” we can say that Hegel has identified the antecedent condition of our phenomenal nature with regard to our responsiveness to the moral law; and that because of this antecedent condition, the moral law, even through a representation of it, cannot be the “angle of inclination for inclinations.”

## *2. Theoretical tensions*

Having considered Hegel’s primary argument against moral purity, I turn to an ancillary

argument that, though overlooked in the secondary literature, provides a compelling reason for why Kant or the Kantian cannot be in earnest about moral purity and which, moreover, directly targets the duty to strive for moral purity. Hegel clearly has Kant and *The Doctrine of Virtue* in mind when he writes:

That consciousness is not in earnest about the perfection of morality is indicated by the fact that consciousness itself shifts it away into *infinity*; i.e., asserts that the perfection is never perfected. What consciousness really holds to be the truth of the matter is only this intermediate state of imperfection, a state nevertheless which at least is supposed to be a *progress towards* perfection. But it cannot be even that; for to advance in morality would really be to move towards its disappearance. That is, the goal would be the nothingness or the abolition, mentioned above, of morality and consciousness itself; but to approach ever nearer to nothingness means to diminish. Besides, 'advancing' as such, like 'diminishing', would assume *quantitative* differences in morality; but there can be no question of these in it. In morality, as in consciousness, for which the moral purpose is *pure* duty, there cannot be any thought at all of difference, least of all the superficial one of quantity, there is only one virtue, only one pure duty, only one morality. (*PH*, §§622-623)

Regarding the shifting position of moral consciousness: there is no disputing that Kant determines a duty of infinite progress toward moral purity without realization, as we have seen in the explication of positive duties above (*MM*, 6:446). Likewise, there is no disputing that Kant has, in the *Groundwork*, defined duty in terms of and made moral worth conditional on *actual* moral perfection (*G*, 4:398, 400; see also *MM*, 6:392-393). These two commitments introduce an evident tension which accords with Hegel's imputation of dissemblance: if we cannot attain moral perfection “in this life” but duty consists in and moral worth is conditional on *actual* moral perfection, then we cannot perform our duty (but only action in accordance with duty) and we cannot be morally worthy. Actual duty and actual moral worth are impossible, existing only as ideals toward which we can strive. Simultaneously we have a duty to strive toward our duty – we have

a particular duty to strive toward duty construed as “the necessity of an action executed from respect for the moral law.”

Regarding Hegel's objections to such striving, he contends, first, that striving for moral purity is “really ... to move towards its disappearance.” Hegel's meaning here is clear: he intends the disappearance of *actual* morality. Since, according to Hegel, actual duty requires empirical motivation, to strive for a moral purity would be to strive for *not* realizing one's duty. And to strive toward not realizing one's duty is to strive toward the abolishment of morality in any meaningful sense. As I am using this ancillary argument to *support* Hegel's contention regarding the necessary incorporation of empirical motives, this objection cannot be included without begging the question.

However, Hegel also contends that a duty of moral progress is “superficial” and fundamentally opposed to Morality – “there can be no question” of it. We may read Hegel's use of “superficial” as designating moral progress as the contrary of true virtue, or the essence or genuine nature [*Wesen*] of Morality.<sup>18</sup> In a subsequent section, Hegel confirms this reading: “But it is the essence of morality to be only the *perfectly pure*; imperfect morality is therefore impure, or is immorality” (*PH*, §626). Said differently, though quantitative progress accords with our common understanding of virtue and self-development, this is only a *semblance* of true virtue, and is opposed to the theoretical commitments of Morality. That is, we think of ourselves and others as improving our moral character through habituation, moral education, and the acquisition of moral integrity. However, this notion of progress conflicts with the strict theoretical

---

<sup>18</sup> For a discussion of *Wesen* see *The Encyclopedia Logic* §112, especially the *Addition*: “When we speak of ‘essence’, we distinguish it from being, i.e., from what is immediate. In comparison with essence, we regard being as a mere *semblance*.”

commitment of Morality according to which we must act exclusively from duty. Hegel's charge invokes these theoretical commitments. Kant has previously and explicitly committed to the mutual exclusivity<sup>19</sup> of duty and pathological motivation: moral purity consists in “the law being by itself alone the incentive, even without the admixture of aims derived from sensibility, and in actions being done not only in conformity with duty but also *from duty*” (6:446; *CPrR*, 5:71; see also Kant's definition of duty in the *Groundwork*, 4:400). Kant has, moreover, insisted that even allowing inclination to function alongside rational motivation is “hazardous” (*CPrR*, 5:72). We might say, then, that Kant, contrary to these commitments, introduces a quantitative notion of virtue. Hegel's response is that this is a category mistake: any degree of virtue simply does not apply to purity of the moral will. But even if one abandons the purity of the moral will, one is making a category mistake in that, insofar as moral purity is put off in the “infinite” and thus unattainable, “progress” is unintelligible. No matter how much progress one makes toward moral purity, one will remain morally impure and will fall short of moral worth.

To illustrate we may consider again Kant's example of a shopkeeper (*G*, 3:397). If we imagine two shopkeepers both of whom require pathological motivation (i.e., motivation grounded in subjective and empirical ends and the corresponding inclinations) to do what duty requires, then neither shopkeeper is due moral esteem. If we then stipulate that one of the shopkeepers needs *less* pathological motivation than his colleague, we may in a commonsense way say this is improvement; but we may not

---

<sup>19</sup> Said differently: The propositions “In performing action X the agent did his duty” and “Pathological motivation was a necessary condition for the agent to perform action X”, where “action X” refers not only to a discrete kind of action but a numerically identical action, are exclusive disjuncts.

attribute moral worth to the shopkeeper who is further along, and thus there has been no advancement *qua* moral worth; both shopkeepers have equally failed to do their duty. More simply, because moral purity can never be actual, moral purity is always infinitely remote. Thus, in a theoretical sense there is no improvement as there is no, and according to Kant can never be, a realization of duty. It is in this sense that Hegel means that “imperfect morality is...impure, or is immorality.”

There is another sense in which striving for moral purity is a category mistake *qua* duty. Moral purity deserves esteem and thus has appeal as a duty – we ought to achieve and be that which is worth of esteem. However, striving for moral purity, when such striving is necessarily Sisyphean, does not have this same appeal. What worth that could ungird duty is there in perpetually falling short of actual moral worth?

Let me amplify this point. So far moral purity as an end *to be achieved* has been considered unchallenged.<sup>20</sup> I considered above how Kant views moral purity as following from the full realization of oneself as an end-in-itself. Moreover, Kant certainly takes it to be essential<sup>21</sup> to the “morality” of an action. Nonetheless, we may concede (I do, though others will not) that only an agent who acts from duty deserves *esteem*, while still doubting whether esteem and moral worth are always overriding reasons especially in light of competing duties. (For a simple example, what if I find it significantly easier to

---

<sup>20</sup> For a particularly cogent challenge to traditional Kantian moral purity, see Herman (1998, esp. 43): “The question comes to this: Must it follow from the fact that the foundations of morality – the determining ground of the will – can only be in the pure practical reason, that (a) the motive of the morally worthy agent must be extramaterial (a miracle every time), or that (b) the desirable *purity* of the motive of duty is in its complete separation from the empirical life of the human agent?”

<sup>21</sup> “What is essential to any moral worth of an action is *that the moral law determine the will immediately*. If the determination of the will takes place *conformably* with the moral law but only by means of a feeling, of whatever kind, that has to be presupposed in order for the law to become a sufficient determining ground of the will, so that the action is not done *for the sake of the law*, then the action will contain *legality* indeed but not *morality*” (CPrR, 5:71).

be generous if I set up a reward system for myself?) This skepticism of moral worth as an overriding end is greatly amplified when moral purity is necessarily unattainable; for now, we must hold *striving* for moral purity and worth to be an overriding duty. To see why this is questionable, consider the other sense of moral perfection that Kant has determined as a positive duty: namely, *comprehensively* achieving one's particular duties. I have the positive duty to achieve, to the best of my ability, all of my duties. The performance of these duties may be done from duty or in accordance with duty – that is, I may do what is particularly required of me from empirical motivation. As we have seen, Kant recognizes that some forms of empirical motivation, namely natural sympathy<sup>22</sup> provoked and cultivated through exposure to the suffering of others, are effective means to performing my particular duty. However, I also have a duty to strive for moral purity that can never be achieved. And this latter duty *may* have lexical priority. These various duties introduce a tension into Kant's theory: If we recognize that empirical motivations may *realize* particular duties effectively, then we may plausibly reduce the effective realization of particular duties by eliminating empirical motivations from our psychological repertoire. In other words, we may plausibly undermine our duty of moral perfection (i.e., comprehensive moral achievement) by pursuing a duty of moral purity, the end of which can never be achieved. Fully realizing perfection may not be compatible with striving for moral purity. If this is the case, then normative force of this latter duty – already in question – is most severely attenuated.

---

<sup>22</sup> Kant distinguishes between free sympathy and natural sympathy and specifically says we should use natural sympathy as a means (*MM*, 6:457).

Lastly, Hegel's objections are compounded by an epistemological assumption; namely, that we can know that we are making progress. This assumption is incompatible with Kant's view of volitional transparency:

The depths of the human heart are unfathomable. Who knows himself well enough to say, when he feels the incentive to fulfill his duty, whether it proceeds entirely from the representation of the law or whether there are not many other sensible impulses contributing to it that look to one's advantage (or to avoiding what is detrimental) and that, in other circumstances, could just as well serve as vice? (*MM*, 6:447; see also *G*, 4:407)

The rejection of this epistemological assumption heightens the first objection regarding the tenability of considering striving for moral purity a duty. If for any given action or instance of motivation I cannot know what is actually motivating me, then I cannot know whether or not I have made progress toward moral purity because I cannot compare a past instance of motivation (about which I have no reliable introspection) with a subsequent instance of motivation about which I am equally blind. If I cannot know whether I have made progress, then I cannot *intelligently* strive toward moral purity.

To see this concretely, we may consider how volitional opacity would again pose problems for the prioritization of duties. Suppose, for example, I judge that I am to a greater degree motivated by duty, but fall short of moral purity in certain scenarios or when confronted by certain countervailing inclinations. Contrary to this judgement (based on faulty introspection and false corroboration from friends), I am in fact rarely and weakly motivated by duty in even ordinary circumstances. Moreover, suppose that, confident in my high degree of moral purity, I am pursuing my indirect duty "to cultivate the compassionate natural (aesthetic) feelings...and to make use of them as so many means to [practical] sympathy" (*MM*, 6:457). By what measure do I judge the success of these practical means, especially in the context of striving for moral purity? Perhaps my



efforts are fruitful, but perhaps not. Perhaps my dutiful activity of cultivating sympathy is having a deleterious effect on my striving for moral purity – specifically, for all I know, my increasing compassion is eroding my capacity to act from duty. In this way, I not only have an internal conflict of positive duties to myself; I am unaware of the conflict and, even if I were to become aware, impotent to intelligently resolve the conflict due to the introspective opacity regarding the effectiveness of my efforts.

To summarize: Hegel argues that Morality is facing a dilemma. On the one hand, Morality requires moral purity, or doing one's duty for the sake of duty. On the other hand, particular duties must be determined – our duty must be actualized. Hegel contends that the actualization of duty requires empirical or pathological motivations. However, Morality cannot incorporate such motivations without violating moral purity. Consequently, the determination of duty remains abstract, and the merely formal determination of duty (respect the right and promote the good) presents Morality with the second horn of the dilemma: the subjective determination of particular duties. This initial argument is unsatisfactory insofar as it appeals to a contentious premise: namely, that the actualization of our duty requires empirical motivation. As such, I supplemented Hegel's original objection by appropriating arguments from the *Phenomenology* which problematize and militate against moral purity.

In the next section I turn to the emptiness objection. In this final extension of the emptiness objection to positive duties derived from the Formula of Humanity, subjectivism again figures centrally.

### III. THE FORMULA OF HUMANITY, POSITIVE DUTIES, AND EMPTINESS

Extending Hegel's emptiness objection to the Formula of Humanity and positive imperfect duties takes two forms: First, I argue that the limitation of negative duties argued for in chapter 4 has the direct corollary of expanding the scope of possible positive duties. Second, I explicate and elaborate on Hegel's argument that the subjectivism of the Moral standpoint and the insufficiency of the Formula of Humanity in determining particular duties allows for the actual elevation of the particular above the universal, which both Hegel and Kant hold to constitute evil.

### **A. The Expansion of Positive Duty**

There is one sense in which the extension of Hegel's emptiness critique is straightforward: the arguments in chapter 4 severely limited the proscription against suicide and deception and, by extension, other putative negative duties. This limitation of negative duties immediately entails the expansion of possible positive duties.

Our performance of positive duties is not unrestricted. Kant is clear that we cannot do an action that is immoral in order to satisfy a positive duty I have to myself or another. This restriction is based on the lexical priority given to perfect duties: perfect duties are narrow, meaning they demand strict adherence, whereas imperfect duties are wide, meaning they may be satisfied according to the agent's discernment (see for example, *MM*, 6:390 and 393). Accordingly, one may never violate a perfect duty, and this includes actions for the sake of positive imperfect duties. This hierarchy and restriction has intuitive appeal: I cannot, for example, steal from one individual (what is rightfully theirs) in order to promote the interests and happiness another.

Perfect duties, and for our purposes perfect negative duties, have significant implications for our performance of positive imperfect duties. This, in turn, means that the arguments undermining the determination of particular duties in chapter 4 changes what particular positive imperfect duties we may have. Simply put, if certain actions that were considered immoral cannot be judged as such by the Formula of Humanity, then those same actions fall into the category of possible positive duties.

The clearest example of this implication is suicide, and more specifically euthanasia. In chapter 4 I argued that either a universal and rational ground for determining good reasons for suicide could not be identified *or* that suicide does not necessitate a disregard for humanity, and therefore an absolute proscription against suicide was not derivable from the Formula of Humanity. It follows, then, that one may commit suicide, including voluntary euthanasia, so long as they have a good reason and do not disregard their own humanity. This has significant and immediate implications for our duty to others: if euthanasia is not absolutely proscribed, then assisting in euthanasia is also not absolutely proscribed; that is, there is not a perfect duty restricting our assisting in voluntary euthanasia. If assisting in euthanasia is sometimes permissible, then assisting in euthanasia may sometimes be *required by duty*. In other words, since euthanasia is not always immoral, euthanasia may then be one's moral end, and we have a duty to promote another's permissible ends. Therefore, euthanasia is not only permissible for others, but also a form of beneficence.

This duty is not entirely unrestricted. That is, the general implication that assisting euthanasia may be a duty does not entail that we must always assist in euthanasia. First, beneficence is a wide duty, and we may determine when and how to promote another's

happiness. This discretion also means that one may withhold beneficence until one determines that the action in question does, in fact, conform to morality. One may, for example, have questions or doubts about an agent's motivation for suicide, the permanence of these motivations, or the relative benefit of different options. However, a Hegelian may concede that an agent does not have to perform every kind or realize every available instance of beneficence; for extending Hegel's critique it is enough to show that actions previously considered impermissible are now compatible with positive duty. In concrete terms, what this means is that it may be our duty – or, at least, one may be performing his duty of beneficence – to help a 24-year-old struggling with severe depression commit suicide.<sup>23</sup>

It is easy to see how this line of reasoning can be extended to other duties. Chapter 2 furnished an important example of lying for the realization of a good end. If an agent were confident that a prosecutor was manipulating the system for pernicious ends and our friend was the victim of this perversion of the otherwise legitimate justice system, we may be obligated (or, more modestly, our actions may be compatible with positive duty) to lie under oath to prevent the miscarriage of justice. This example, however, may be seen as an extreme exception; and while such exceptions serve a purpose in chapter 2, our current goal is to show a general expansion of possible duties.

---

<sup>23</sup> I have drawn this example from a real case of a Belgian woman who suffered (and may suffer) from severe chronic depression and, after extensive analysis by different medical professionals and following the legal requirements for voluntary euthanasia in Belgium, was granted the right of voluntary euthanasia even though she was, at the time, only twenty-four and physically healthy. She eventually decided against euthanasia. "24 & Ready to Die," a short documentary of her story by *The Economist*, is available here: <https://www.youtube.com/watch?v=SWWkUzkfJ4M> (accessed March 22, 2017).

The arguments in chapter 4 accord with this objective. We saw in chapter 4 that deception may be permissible in some cases given the right social conditions or normative expectations. This exceptionality was possible because deception is considered wrong, according to one line of Kantian reasoning, because it circumvents or coopts another agent's rational will; however, such circumvention is not always impermissible and is, indeed, sometimes a necessary and benign means to a good end. In chapter 4 these arguments were developed to show that deception (or other forms of untruthfulness) are often not violations of perfect duty. However, if this is true, then deception (and other forms of untruthfulness) does not constitute a moral obstacle to positive duty. Therefore, one may deceive or tell falsehoods as means of beneficence and one's own self-perfection.

Consider, for example, a modification of an example from chapter 4: You and a friend are at a work-related party and you observe the possibility of an extremely embarrassing scenario. Perhaps your friend's ex is in the other room, and you rightly discern that this could lead to an explosively inappropriate interaction. You also know that such an interaction would be damaging to your friend's prospects at work which he greatly values and is genuinely good for him. You consider the options and decide that the best available option is to concoct a story that gets both of you to leave. Your deception genuinely promotes your friend's happiness, and, we may stipulate, your friend is grateful for the intervention and cooption of his will (though most would not express it in the latter terms). As in the case of euthanasia, such a duty is wide: we cannot demand that you lie to your friend, and we certainly cannot compel you to do so; nonetheless, you

could lie to your friend and this deception could be compatible with genuine and permissible beneficence.

This example of deceiving a friend is of particular interest because, as in the cases from chapter 4, the permissibility of the deception goes beyond consent. One cannot deceive if one has immediate permission to deceive – as Korsgaard notes, one cannot take what is freely given. In this case, the lying friend deceived permissibly not because he had explicit permission (we can imagine that the issue had never come up – he was not following an agreed upon protocol), but rather because he possessed a historically and culturally contingent understanding of friendship and what friendship allows and, indeed, requires. In other words, the agent acted according with normative expectations.

As with the arguments in chapter 4, these examples do not merely demonstrate that there are rare exceptions to a generally applicable rule. Rather, the problem with the derivation in the previous chapter was that no consistent and independently tenable principle could be identified in Kant or his defenders that could serve as an intermediate premise for the derivation of negative duties from the Formula of Humanity. As such, the current objection is not that there may be exceptional cases where one may, for example, deceive as a means to realizing a positive duty; but rather that, in general, Morality does not provide a principled or formal exclusion of such actions.

## **B. The Manifold of Duties, Conflict, and Subjectivism**

In section II I elaborated on Hegel's subjectivism objection clarifying why Hegel contends that Morality necessarily faces and must resolve the problem of subjectivism. This problem has two different aspects. First, Hegel argues that Morality leaves it up to

the subject to determine the quality or character of an action; that is, in Hegel's terms, the subject determines what is essentially good in an action. As has been argued in previous chapters, this subjective determination of what is essential to an action has severe implications for the generation of a contradiction when universalizing the maxims of natural actions (see chapter 2, 57ff). However, in considering the subjectivism objection above, and specifically Hegel's presentation of it in *Faith and Knowledge*, another aspect of this objection was disclosed: namely, that if Morality does not determine a context specific duty, then it is up to the subject to do so. In other words, the subject, without concrete guidance from Morality, chooses his particular duty from a manifold of duties. In this way, the intersection of the subjectivism objection and emptiness objection can clearly be seen: the application of the Formula of Humanity is dependent on the subjective determination particular concrete duties, and therefore, insufficient for determining our actual duty.

Explicating this objection has several components: First, I reconsider Hegel's contention that there is a "manifold of duties" and a choice is required to determine our actual duty and situate this contention within the context of a conflict of duties. Specifically, Kant contends that on conceptual grounds there can be no conflict of duties, and so, initially, Hegel's challenge appears to be mistaken. However, with regard to positive duties, Kant does recognize competing "grounds of obligation" and that the stronger ground of obligation must prevail over other competing grounds. As such and in accord with Hegel's objection, which duty is our actual or "valid" duty remains undetermined by Morality as such. Hegel's radical contention is that this indeterminacy

of actual duty creates a theoretical space in Morality for the capacity for evil in the form of elevating the particular above the universal.

In the following section I elaborate on Hegel's objection by considering and modifying a famous example from the canon of moral philosophy – namely, Judith Jarvis Thomson's Henry Fonda analogy. I argue that this analogy represents a case where there is a "threshold of opprobrium" but that Morality lacks the theoretical resources for identifying when this threshold has been violated.

In the last section I consider objections to Hegel's objection. Chief among these objections is a rebuttal based in a rigoristic reading of Kant and our imperfect duties. This reading of Kant is represented by Peter Atterton (2007) who argues for the lexical priority (within the context of imperfect duties) of moral self-perfection and beneficence. Accordingly, if this lexical priority were tenable, Hegel's objection would be significantly attenuated. I reject a rigoristic reading of Kant and argue that we have a strict duty to ensure our own happiness and an independent duty to pursue our natural perfection. Moreover, our duty of beneficence is compatible with *privileging in action* those who are closer to us, including ourselves. Accordingly, the full strength of Hegel's objection remains intact.

### *1. Clarifying Hegel's objection in the context of conflicting "grounds of obligation"*

Here again is the relevant selection from *Faith and Knowledge*:

As was shown earlier, the duties and laws in this system [of formal morality] make up an infinite, dispersed manifold every element of which is equally absolute. So the manifold makes a choice necessary. This choice [between duties] is strictly a subjective matter, for the objective aspect, the form of universality, is what is common to all [duties equally]. Now, we cannot think up any actual case of an acting that would not have several sides; for every intuition of an actual case



is infinitely determinable through the concept [i.e., abstract duty]. Some of these sides must constitute valid duties so that in obeying some duties, other duties will be violated and in violating some duties other duties will be obeyed. (GW, 184)

Hegel does not qualify this objection. That is, he appears to consider every particular duty, regardless of kind, to have the same moral weight. This, of course, directly opposes Kant's view that perfect duties out-rank and supersede imperfect duties. Therefore, according to Kant, there cannot be a genuine conflict of perfect and imperfect duty. This does not, however, rule out the possibility of there being a "moral residue" of regret or guilt (see Timmerman, 2013). Nonetheless, I wish to put this form of conflict to the side and focus, rather, on the possibility of conflict between positive duties. This circumscription allows for illustrating the kind of objection Hegel has in mind – the insufficiency of formal morality to determine particular duties and the essential role of *Sittlichkeit*.

First, let us get clear on what Kant says about the conflict of duties and, specifically, the conflict of competing positive duties. In the *Introduction to the metaphysics of morals* Kant presents the "preliminary concepts" that will be of concern throughout the text. In this conceptual taxonomy, he directly addresses the possibility of duties conflicting. His position, broadly, is that a conflict of objective duties is impossible, but a conflict may obtain between subjective "grounds of obligation." Concerning the former impossibility, Kant writes:

A conflict of duties (*collisio officiourum s. obligationum*) would be a relation between them in which one of them would cancel the other (wholly or in part). – But since duty and obligation are concepts that express the objective *necessity* of certain actions and two rules opposed to each other cannot be necessary at the same time, if it is a duty to act in accordance with one rule, to act in accordance with the opposite rule is not a duty but even contrary to duty; so a collision of duties and obligations is inconceivable (*obligationes non colliduntur*). (MM, 6:224)

Kant here is arguing from a conceptual necessity, and much depends on the terminology used by Kant. Kant justifies the impossibility of conflicting duties by arguing that “two rules cannot be necessary at the same time.” Following Timmermann (2013) “rules” in this context signifies “*general* commands or duty *types*.” As a matter of internal consistency, such rules cannot both be necessary at the same time – if one were necessary, then the other could not be possible. Timmermann also thinks the same line of reasoning can be used at the token level; that is, regarding obligations, two obligations cannot be simultaneously practically necessary.

However, the conceptual exclusion of a conflict between what is practically necessary does not exclude all conflict. Kant continues in the same passage and contends that there can be a conflict between the *grounds* of obligation, though not obligations *per se*:

However, a subject may have, in a rule he prescribes to himself, two *grounds* of obligation (*rationes obligandi*), one or the other of which is not sufficient to put him under obligation (*rationes obligandi non obligantes*), so that one of them is not a duty. –When two such grounds conflict with each other, practical philosophy says, not that the stronger obligation takes precedence (*fortior obligatio vincit*) but that the stronger *ground of obligation* prevails (*fortior obligandi ratio vincit*). (MM, 6:224)

According to Timmermann, “ground of obligation” must be restricted to the category of wide duties (see pages 43-50, and especially n. 23). Drawing upon various texts for corroboration, Timmermann argues that Kant distinguishes between two different “laws”: strict and wide laws which, in the *Metaphysics of Morals*, correspond to duties of right and duties of ethics, the latter of which are defeasible. The latter also provide only grounds of obligation and not strict, that is, necessary obligation. In other words, we have a reason to act but are not practically necessitated to act. This is because, as has been

discussed above, ethical duties are both conditional on being permissible (i.e., they cannot violate negative duties) and contingent (i.e., they only become relevant under certain conditions). As such, two grounds of obligation can conflict.

This still leaves “grounds of obligation” unclear – in other words, we still need an answer to the question “Between what is there a conflict?” A ground of obligation is the reason why a principle becomes obligatory within a particular situation or context. “In particular, a ground of obligation arises when an agent correctly applies an ethical principle to a concrete case” (Timmerman, 2013, 48). Take for example genuine need: a general principle of beneficence is furnished by the Formula of Humanity, and a particular individual may be in a particular state of affairs contrary to their good. Presented with our principle and the facts of the matter we may still pause to consider whether beneficence on our part – intervening in the individual’s state of affairs to promote his good – applies. If we correctly apply a principle to the facts we generate a ground of obligation, or reason for why we should act.

With this restriction and clarification, we can see how grounds of obligation can conflict. Timmermann furnishes an example (originally given by Vigilantius in his notes to Kant’s lectures on the *Metaphysics of Morals*). A conflict could hold between our general duty of beneficence and our duty to care for our family (Timmermann, 2013, 47). In this case, there is good reason for both actions, though failure to act would not violate right – that is, both duties are ethical. Also, there is no theoretical conflict: each duty is a product of compatible laws. The conflict arises when, in trying to apply the laws, the means are not available to achieve both. However, since neither is an *obligation* in the

strict sense (a strict or narrow duty) there is no conflict of obligation, only a conflict of reasons.

How can such conflicts between grounds of obligation be resolved? Kant's pithy answer is that the stronger ground of obligation prevails. If we have a stronger reason for action, or as Timmermann notes, if we have two reasons for action, the greater "weight" of reasons determines what we ought to do. In the above example duty to one's parents is "declared" to have the stronger ground.

This is an unsatisfactory answer. However, Kant contends that we cannot reach any greater precision in resolving such conflict. In the Introduction Kant outlines the need for a *Doctrine of Virtue* to include a Doctrine of Method:

The reason [for a Doctrine of Method] is that the doctrine of right has to do only with narrow duties, whereas ethics has to do with wide duties. . . . But ethics, because of the latitude it allows in its imperfect duties, unavoidably leads to questions that call upon judgment to decide how a maxim is to be applied in particular cases, and indeed in such a way that judgment provides another (subordinate) maxim. (*MM*, 6:411)

This necessary role of judgment leads Morality into the domain of *casuistry*. However, casuistry, as understood by Kant, is a disappointing resource for the kind of moral judgment required. Kant thinks that casuistry can be neither systematic nor scientific (which would reduce it to "dogmatism") and considers it a mere "scholia" – or marginal note – appended to a doctrine of ethics. This minor role of casuistry is borne out in Kant's development of casuistry in the *Doctrine of Method*: regarding duties to oneself and others, the Doctrine is divided into two sections (with a third section dedicated to duties to God). The first section concerns teaching ethics and the second concerns ethical ascetics. Both sections focus on virtue construed as giving lexical priority and motivational exclusivity to duty. That is, the method concerns how to teach the proper

understanding of duty in relation to inclination and happiness, and how to cultivate both a Stoic and Epicurean attitude to misfortune and duty (we should be resolved to do our duty in the face of countervailing reasons, and cheerful in our duty regardless of unhappiness disproportionate to our virtue). Only very briefly does Kant consider the substantive resolution of problem cases, or judgment of *content*, when he writes that students (children) should be given the opportunity to solve problem cases on their own; even here, however, no grounds are provided for resolving problem cases.

Timmermann's conclusion about Kant appears to be unavoidable: with regard to positive duties such as beneficence, we can go no further than determining our general duty; beyond that "Kant trusts that a virtuous person will reach the right conclusion" (Timmermann, 2013, 55). This is already enough to establish, in a limited sense, the emptiness of the Formula of Humanity: If relying on virtue means that we must rely on the agent who puts duty first and is motivated by duty, then this is tantamount to saying that virtue is not sufficient for determining content since duty, in the abstract, has already been shown to be insufficient for determining my particular duty. Prioritizing duty and acting from duty cannot aid in determining *what one's duty is*. If "virtue" means something more substantive – such as a practical knowledge of what is required in a particular, socially and culturally contingent context – then this is tantamount to saying that Morality is empty insofar as it requires knowledge of an immanent normative content that is not furnished by abstract Morality. Hegel's objection, however, is even more severe.

It may appear that Kant's and Hegel's views (at least as they pertain to positive duty) accord: ultimately the subject must judge what his particular duty is. The difference

between Kant's and Hegel's views – or, in what the extension of Hegel's objection consists – lies in the severity of what is entailed by the absence of a systematic resolution of problem cases. Kant does not discuss casuistical cases to any great degree; moreover, the three cases he does provide when discussing beneficence all represent marginal examples of conflict.<sup>24</sup> Thus, Kant (and Kantians) hold that leaving the resolution of difficult cases up to the judgment of the virtuous agent entails only rare exceptions to a generally applicable theory (see, e.g., Wood 2008, chap. 2) Hegel, to the contrary, contends that the subjectivizing of moral judgment entails the capacity for *evil*.

Where all previously valid determinations have vanished and the will is in a state of pure inwardness, the self-consciousness is capable of making into its principle either the *universal in and for itself*, or the *arbitrariness of its own particularity*, giving the latter precedence over the universal and realizing it through actions – i.e., it is capable of being *evil*. (PR, §139)

The “universal in and for itself” Hegel has identified in the preceding section – it is abstract duty, the *right* and promotion of the *good* from the standpoint of Morality, which is to say, from a standpoint where all right and every rational and free subject's end is equally considered. Or in Kant's terminology, where *humanity* is the incomparable locus of moral concern and the ground of duty. The Formula of Humanity requires that we act for this universal end. However, Hegel says that a subject, in determining his own particular duties, may act for his “own particularity.” What Hegel means is that an agent

---

<sup>24</sup> Kant asks to what extent one has a duty to be beneficent, and answers with confidence that one should not be so generous so as to become in need of beneficence himself. As we shall see below, there is textual support for the view that Kant holds all beneficence up to this point to be required by duty. Likewise, he asks whether a lord or master who robs a servant or serf of his freedom but makes the same abundant in happiness may be called beneficent. While I do not declare such cases to be entirely free from conflict, they also do not represent counter-intuitive cases for modern readers. Most readers will recognize self-improvement as a good reason to refrain from continued generosity; and many will likewise see freedom and self-determination as overriding values relative to happiness. Such problem cases, that is, do not represent a problematic emptiness of Morality.

may elevate his own interests and his own happiness above the universal, and therefore above duty abstractly understood. Of course, an agent may not do this *explicitly*; an agent may not make his own interest an abstract law contrary to the Formula of Humanity; an agent may not, for example, make egoism his governing law. What the agent may do is determine what, in concrete terms, is required by the abstract principle of Morality and then judge an ancillary principle which promotes one's own particularity as compatible with or *required by* the law of Morality, act on this ancillary principle, and thereby realize self-interest in practice without violating abstract duty. Simply, one's own concrete particularity is compatible with abstract duty, with the Formula of Humanity. This is, on Kant's own terms, constitutive of evil: the practical elevation – the making exceptional – of the individual (see Dietrichson, 1969; Korsgaard, 1996, 92).

We have, then, come full circle: in chapter 1 I addressed the Kantian interpretation of the objective of the Formula of Universal law – to proscribe maxims which are exceptional in character. I argued that this abstract proscription, first, does not progress beyond the Formula of Universal Law itself; second, this abstract proscription does not effectively proscribe concrete actions contrary to abstract duty. The same general argument is made in chapter 4: the Formula of Humanity requires independent moral content or dubious moral judgment to derive particular duty. Here, the absence of particular judgment, and the subsequent subjectivizing of judgment, allows for the actualization of the precise evil Morality condemns.

## *2. Elaborating Hegel's objection*

Should we accept Hegel's contention? Hegel does not adequately elaborate this consequential objection with examples, especially as it pertains to positive duties.<sup>25</sup> I propose illustrating how compelling Hegel's contention is by considering and modifying one of the most influential examples from the canon of moral philosophy, namely, Thomson's (1971) Henry Fonda analogy. Simply, Hegel's challenge may be presented as a question Morality cannot sufficiently answer: Appealing to nothing else than the Formula of Humanity, at what point *must* Henry Fonda act beneficently and heal the terminally ill stranger?

While not Thomson's most famous example – this honor must go to the violinist analogy<sup>26</sup> – the Henry Fonda example has more direct relevance. The violinist example

---

<sup>25</sup> Hegel's elaboration of his objection with examples in the *Philosophy of Right* is as follows: "Theft, cowardice, murder, etc. as actions – i.e. as products in general of a subjective will – have the immediate determination of being the *satisfaction* of such a will, and hence of being something *positive*, and in order to make the action into a good one, it is merely a question of knowing this positive aspect as my *intention* in performing the action, and this aspect is then the *essential* factor in determining the action as good, because I know it as the good in my intention. Theft in order to benefit the poor, theft of desertion in battle for the sake of one's duty to care for one's life or one's (perhaps even impoverished) family, murder for hatred and revenge...all of these deeds, by virtue of the positive aspect of their content, are in this way transformed into well-intentioned and consequently good actions." (PR, §140d). Each of these examples involves a violation of negative duty for the sake of some good end. The possibility of violating putative negative duties by changing the intentional content was taken up extensively in chapter 2 (pgs. 56ff). Here we are concerned with a conflict of grounds of obligation and the subjective determination of one's actual concrete positive duty.

<sup>26</sup> The violinist example does, I think, have indirect relevance and may be helpful in illustrating the severity of Hegel's objection by simply representing how fundamentally opposed moral judgment can be. To briefly review the violinist analogy, Thomson asks the reader to consider the following scenario: You wake up to find that a society of music lovers has connected a famous violinist, who is dying of kidney disease, to you so that your kidneys will function like a living dialysis machine. It is explained to you that only your kidneys are a match for this purpose, and that unplugging the violinist will result in his death. The violinist requires (in the original presentation) remaining attached for nine months (to represent a pregnancy), after which you may detach yourself from him and be free of any further responsibility. Also (in the original presentation) no harm to you will result from the use of your body.

The influence of the violinist example makes it particularly helpful for illustrating the possibility of radical disagreement about moral obligation and the subjectivism Hegel contends is inextricable from moral judgment in the context of formal Morality. The analogy has engendered interest in and debate on the overt intuitional grounds undergirding Thomson's argument. Concerning the permissibility of severing oneself from the famous violinist, Thomson writes: "If anything in the world is true, it is that you do not commit murder, you do not do what is impermissible, if you reach



involves taking action which will lead to the death of another, and is designed to confront a recognized right to life and one's own legitimate interests. Indeed, Thomson explicitly develops the example to demonstrate that no injustice is perpetrated by the individual who severs himself; thus, it concerns what Kant terms negative and strict duty. While the Henry Fonda example, and its subsequent modification, has the same broad desideratum, it is also designed to show that *refraining* from an action that would save a life also does not constitute an injustice. While Thomson's objective may be to show that (what he takes to be) extreme positive action is not necessitated by another's right to life, the example may readily be appropriated to show Hegel's challenge to Morality; namely, the

---

around to your back and unplug yourself from that violinist to save your life" (1971, 51). If we take Thomson literally, then Thomson finds this moral judgment, though clearly based on her own judgment of what is true, to be undeniable. And yet there is empirical evidence of significant disagreement. Schroer and Bradner (2011) have investigated the influence of personal relation on individual's intuitions when presented with the violinist example (specifically, they investigate what intuitional shifts occur if the violinist is identified as one's half-sibling). Though not directly investigating Thomson's putatively undeniable claim, in order to investigate the influence of personal relation on moral obligation in the violinist scenario, a baseline was needed; that is, how individuals respond to and judge their moral obligation vis-à-vis the original scenario had to be established. Perhaps surprisingly for some, only 34% judged that they were "Definitely Not" morally obliged to stay connected to the violinist. Interestingly, 12.9% answered "Definitely Yes" to the same question, and an additional 18.3% answered "Probably Yes" – that is, 31.2% judged there to be a probable or definite moral obligation to remain connected to the violinist stranger.

Importantly, there is no claim being made here – nor by the authors, I hasten to emphasize – that such intuitions have theoretical bearing on our actual moral obligation. Rather, my only present interest is the variability of conviction, and conviction of the greatest strength. For Thomson, unplugging oneself from the violinist represents the absolute extreme pole of what is permissible, and yet this judgment is not shared. As we have seen, some hold to be impermissible what is supposed to be absolutely permissible. What this illustrates is the extreme variability of moral judgment concerning the conformity of an action to duty; said differently, our judgment of when we pass the threshold of moral obligation is subjective. Furthermore, this radical disagreement poses the question: If there is such actual variability, what other possible variability could there be and what are the grounds for such variability? Hegel's general contention is that within Morality the sole ground for judging the conformity or nonconformity of an act to duty is the subject either because the subject determines the intention, and thus essence, of the action or (as we shall see below) the subject determines the relative priority of the action. A subject's judgment alone, however, is arbitrary. More specifically, and with particular relevance for positive duties, Hegel contends that actual particularity may be elevated above universality in a way that conforms with abstract universality. That is, one may judge the opprobrious to be compatible with duty.

absence of formal criteria for determining what is required by positive duty and the corollary capacity to do evil compatible with abstract duty.

The analogy is simple and brief: Thomson asks the reader to consider a situation where she is terminally ill on the east coast and only “the touch of Henry Fonda’s cool hand on her fevered brow” who is on the west coast can heal and save her (Thomson, 1971, 55). Thomson contends that while it would be “frightfully nice” (I take this to mean supererogatory) for Fonda to travel to the east coast to heal her, it is not a violation of her rights if Fonda does not. Thomson expresses this as an agent having no right against Fonda that he should do this. We may transpose this into Kant’s terminology and say that Fonda does not have an *obligation* to heal the ill agent; or, that it is not *practically necessary* for him to do so; there could be reasons other than a violation of negative duty for why Fonda is not required to perform this act of beneficence. Though importantly, for Kant and Kantians there must be good reasons for Fonda to refrain. Later in the article Thomson reprises and modifies the scenario such that Henry Fonda need only walk across the room and touch her forehead to save her life. Thomson still insists that the agent has no *right* against Fonda; but he does hold that Fonda “ought” to save the agent’s life, and that it would be “indecent” of him not to.

I have no quarrel here with Thomson’s contention that the agent has no right against Fonda, or, alternatively, that Fonda – in Kant’s technical sense of a strict duty – does not have an obligation to help. I do, however, contest Thomson characterizing Fonda’s failure to act in the modified scenario as “indecent” and that such indecency is what grounds the ought. I also dispute categorizing flying to the east coast as

supererogatory: failing to act in either a case I judge to be egregiously wrong, opprobrious, and condemnable.

My disputes are based on mere intuition. However, for my purposes I only need assume is that there is a threshold beyond which refusal to help would be a violation of the duty of practical love. Where this threshold lies, I do not determine, as this absence of particular determination is what is at issue. What I am proposing, then, is that we imagine a spectrum of actions and a threshold beyond which failure to act constitutes violations of positive duty.<sup>27</sup> We may interpret Hegel as contending that there are no formal grounds furnished by the Formula of Humanity for determining when this threshold has been violated, and, thus, the subject alone is the judge of which determinate actions constitute or violate duty. The reason there are no formal grounds for determining the violation of the threshold is because there is a “manifold of duties” – or, in Kant’s stricter terminology, a diversity of possible duties and, consequently, competing grounds of obligation – and moral judgment is required for ranking and prioritizing these duties; or, in Hegel’s terminology, it is left to the subject “to discover...a positive aspect in every action and hence a good reason and intention underlying it” (*PR*, §140d) and to rank each action relative to others. In simple terms, Henry Fonda may have subjectively determined *good reasons* for not healing a stranger, and these good reasons may formally conform to duty. In this way, the Henry Fonda analogy illustrates the theoretical space in Kant’s theory for the capacity to do evil which conforms to abstract duty: by appealing to

---

<sup>27</sup> The opposite is also true: there is a threshold where helping, or continuing to help, would violate duty, and specifically duty to oneself (one could, for example, dedicate oneself so entirely to the healing of others such that one becomes nothing but a means to the happiness of others).

subjectively determined good reasons, a subject may act for his own particularity in violation of universal demands. To do so constitutes evil for both Hegel and Kant.

What kind of judgment could this be that would, in the mind of the subject, justify refraining from a given act of beneficence? Many, I suggest. Fonda may see his duty lying with his family, and may see his potential act of beneficence as undermining these duties which have greater weight. So, for example, if traveling across the country to heal a stranger is judged as “frightfully nice” or supererogatory, then a moderate duty to oneself or one’s family (e.g., a business appointment or an anniversary) may override this duty. Similarly, Fonda may see duties he owes to himself as overriding this particular duty of beneficence, and he may even see the realization of these former duties as constituting the realization of greater duties to others. So, for example, an agent may see his realization of a duty of self-perfection or ensuring his happiness as having greater weight *in itself* or as contributing, in a more substantive and extensive way, to beneficence. Additionally, one may see a particular act of beneficence as violating stronger grounds of obligations for beneficence; so, for example, Fonda might judge his traveling across the country to be presumptuous, a violation of privacy, embarrassing, etc. Indeed, reasons for not doing our positive (and negative) duties are various, common, and banal. According to Hegel, in his subsequent discussion to the passage above, when we fail to do our duty when we should it is not because we *adopt* evil, but rather because we adopt what we falsely judge to be the good: “Thus it has been said that there is in fact no such thing as an *evil man*, for no one wills evils for the sake of evil – i.e. the *purely negative* as such – but always something *positive*, and hence...always something good. In

this abstract good, the distinction between *good* and *evil*, as well as all actual duties, has vanished” (*PR*, §140d).

In sum, Fonda determines when and if he crosses the threshold of impermissibility and even the threshold of opprobrium, and this choice is arbitrary: it is not based on a rationally defensible derivation of duty from the Formula of Humanity but on his own judgment. As we have seen, only general positive duties are derivable from the Formula of Humanity. One might respond that Fonda’s particular duty is obvious or self-evident. The untenability of such a claim is evident, I think, in the already radical divide between actual judgments of duty in the scenario: Thomson thinks flying across the country to heal the terminally ill would fall into the category of “frightfully nice” – the supererogatory – and not crossing the room to do so is “indecent”; I, on the other hand, judge that not doing either action (barring legitimate opposing claims of duty) is genuinely condemnable. This disparity between judgments is also represented by the radical disagreement over the violinist example (see footnote 26).

The consequences of Hegel’s objection should not be understated. There may be actions for which one cannot judge that it, as a duty, is overridden by competing grounds of obligation. That is, in Hegel’s terminology, it may be difficult to see how refraining from some actions, or choosing a competing ground over another, could be subsumed under the *good* (*PR*, §140d). Walking across the room to heal another is one; pulling an innocent child out of a puddle is another. However, these are psychological and empirical limitations and not necessary limitations imposed by reason and Morality. What the preceding objection entails is that *formally* the Formula of Humanity is compatible with

not acting in these cases and that the subject's own judgment is a necessary condition for the dutifulness of the action.

### 3. *Objections*

A few objections require our consideration. I will first briefly consider the objection that Hegel has shifted our moral deliberations and determinations from maxims to actions, and thus misunderstands Kant. A second objection requires a more extended analysis.

One may object that because we do not have a duty to promote our own happiness, the resolution to the question of when and how much we ought to be beneficent is clear: we ought to be beneficent up until the threshold where more beneficence would cause more harm than good. And any duties to ourselves (duties of natural and moral self-perfection) must be understood in terms of constituting means to the practical love of others.

One may object that Hegel has shifted the locus of moral investigation from an evaluation of maxims and moral worth to particular duties understood in terms of actions. This is an objection I have considered in previous chapters (see chapter 1, 19f and chapter 2, 65f); as such I will only briefly review the responses developed there. First, Hegel's concern for actions – and the possibility of opprobrious actions being compatible with formal duty – does constitute a transition, but also an external critique of Kantian Morality. The crux of Hegel's contention is simple: There is a point at which the failure to *act* is evil; and because Kant leaves the determination of when acting is dutiful up to the agent, he fails to provide the theoretical grounds for avoiding this evil. Furthermore, this deficiency of Morality is a genuine concern within the Kantian literature (see Dietrichson, 1969; Korsgaard, 1996); indeed, as we have seen (chapter 2, 54f ) a parallel

concern, with regard for the Formula of Universal law, underlies the “problem of natural actions” and Korsgaard’s extensive rebuttal. In short, the fact that the moral law, in this case the Formula of Humanity, is compatible with opprobrious and condemnable actions represents a fundamental failure of Morality.

A second objection requires more consideration. The extension of the emptiness objection I have been considering has two necessary conditions: first, that there is a manifold of competing duties, or “grounds of obligation,” and two that the subject determines the relative weight, or priority, of these grounds. There is a rigoristic reading of Kant that threatens both of these conditions. On a rigoristic reading of Kant (and specifically the *Metaphysics of Morals*) moral self-perfection and beneficence – that is, promoting the happiness of others – are prioritized such that all other duties, both direct and indirect duties to oneself, are superseded by them. Such a prioritization, while not completely undermining Hegel’s objection, would significantly circumscribe the “manifold of duties”: we would not have competing grounds of obligation directly related to our own happiness or natural perfection, and thus, we could not rank a duty to ourselves as outweighing a duty of beneficence. As such, elevating our own particularity above universality would be significantly restricted.

Before turning to an explication of the objection, the limits of this objection must be acknowledged. Even if the manifold of duties were circumscribed, and moral perfection and beneficence were given lexical priority, it would still be up to the subject to determine to whom he is beneficent, and what weight is given to various forms of beneficence. This subjective determination may allow for the elevation of the particular

over the universal. Still, the objection represents a step toward undermining the radical extent of Hegel's objection.

Peter Atherton (2007) provides such an interpretation of Kant, and argues that a rigoristic reading of Kant is compatible with the following principle of beneficence, which he labels "R":

R      If we are able to do an act of the kind prescribed by the imperfect duty of beneficence, then we ought, morally, to do it unless, in the same circumstances, we perform some other act of the kind prescribed by imperfect (or perfect) duties. (136)

Atherton supports this reading with a key passage from the *Metaphysics of Morals* which states that "a wide duty is not to be taken as permission to make exceptions to the maxim of actions, but only as permission to limit one maxim of duty by another (e. g., love of one's neighbor in general by love of one's parents), by which in fact the field for the practice of virtue is widened" (*MM*, 6:390). In sum, then, one may not permissibly refrain from one's duty for arbitrary reasons – that is, for one's own inclination or subjective interest. Rather, one must have a good reason, or ground of obligation, which justifies refraining.

On its own, R does not constitute a challenge to Hegel's objection. Hegel, it must be remembered, has not argued that an agent may *arbitrarily* forsake duty; but rather, that determination of what is or conforms to one's particular duty is arbitrary.<sup>28</sup> However, Atherton further argues for a particular relation between positive imperfect duties, including the lexical priority of moral self-perfection and beneficence.

---

<sup>28</sup> "The self-consciousness knows how to discover a *positive* aspect in its own end; for this end, as part of the purpose of an *actual concrete* action, necessarily has a positive aspect. By virtue of this positive aspect, [which it regards] as a *duty and admirable intention*, the self-consciousness is able to assert that its action is good both *for others* and *for itself*." (*PR*, §140)



Crucial to Atterton's argument is a passage from the *Metaphysics of Morals* where Kant specifically considers the latitude, and thus possible limitations on, the positive duty of beneficence:

for if the law can prescribe the maxim of actions, not actions themselves, this is a sign that it leaves a latitude (*latitudo*) for free choice in following (complying with) the law, that is, that the law cannot specify precisely in what way one is to act and how much one is to do by the action for an end that is also a duty. – But a wide duty is not to be taken as permission to make exceptions to the maxim of actions, but only as permission to limit one maxim of duty by another (e. g., love of one's neighbor in general by love of one's parents), by which in fact the field for the practice of virtue is widened. (*MM*, 6:390)

He refers to this passage and the principle expressed in it as “P”: making exceptions to maxims of positive duty, including the maxim of beneficence, is impermissible. A maxim of positive duty may only be limited by the fulfillment of another maxim of positive duty (Atterton, 2007, 142).

With this principle (P) established, two other premises are central to Atterton's argument. First, he argues that moral self-perfection is conditional on doing virtuous acts, such that one must make sacrifices of beneficence in order to develop full virtue. Second, he argues that our happiness cannot be considered a direct duty, but only an indirect duty, in that if we do not maintain our happiness we may be tempted away from our duties of moral self-perfection and beneficence. From this he concludes that particular acts of beneficence are limited only by the general duty of beneficence itself. Our own happiness is subsidiary to our other imperfect duties: “Happiness is a subsidiary end that I adopt in order to fulfill what is really my (direct) duty, which in this case is self-perfection” (Atterton, 2007, 153). And since self-perfection requires sacrificial beneficence, our own happiness is subsidiary to promoting the happiness of others. Atterton's second argument is of particular importance, since without this limitation my

own interests and inclinations could justify exceptions to beneficence: “If my own happiness were genuinely a duty in its own right, then P would permit exceptions to the fulfillment of wide duties like beneficence in the interests of inclination directly, the effect of which would be to increase the latitude for the performance of beneficence enormously” (*ibid.*). That is, if my own happiness were a positive duty, then limiting my maxim of beneficence by a maxim of self-love would be compatible with P.

It is only Atterton's second argument that is of present concern – that virtue requires sacrificial acts of beneficence does not threaten Hegel's subjectivism objection. The argument, I believe, is mistaken on several fronts. First, Atterton deflates natural perfection to moral perfection. Second, the argument that our own happiness is only an indirect duty and is subsidiary to moral perfection and beneficence is mistaken. The crux of Atterton's error, I believe, is an elision between *promoting* my own happiness and *ensuring* my own happiness: we have a direct duty, albeit a *negative* duty, to ensure equal consideration of our own happiness. Third, Atterton does not attend to the relative weight, evinced in Kant's own example of maxims of “love of one's neighbor” and “love of one's parents,” which may be given to competing moral ends and is subjectively determined. Consequently, Kant's limiting condition – that one maxim of duty can only be limited by another – does permit exceptions to beneficence to be made based in my own inclinations and interests.

First, let us briefly consider natural perfection. Atterton rightly observes that in the *Metaphysics of Morals* the duty of natural perfection “is conjoined with, and also indispensable to, the duty to perfect oneself morally” (Atterton, 2007, 150). From this he appears to conclude – since there is no further consideration of a duty of natural

perfection *per se* – that natural perfection is subsidiary to moral perfection. This false conclusion is drawn, it appears, from the correct observation that the highest category of natural perfection pertains to understanding and specifically “of those concepts that have to do with duty” (*MM*, 6:387). This priority, I think, is indisputable in Kant. However, our duty of natural perfection is not limited to the cultivation of the understanding; in his further explication of this duty (*MM*, 6:444-446) Kant clearly includes the cultivation of all our natural powers (powers of spirit, mind, and *body*) within the duty of natural perfection “as a means to all sorts of possible ends” (*MM*, 6:444). The cultivation of these powers is included as a duty because we must be “in a pragmatic respect a human being equal to the end” of our existence. That is, we must – to the extent possible for each of us – cultivate the capacity to realize the various ends set by our humanity.

This inclusion of natural perfection *per se* as a duty means that maxims of natural perfection (maxims, for example, of “going for a run” or “going to the museum” or “taking an art class” or “going backpacking”) can permissibly limit my maxims of beneficence just as a maxim to “love one’s parents” does. Moreover, just as in Kant’s example (and as shall be corroborated below) the subject may – according to his own judgment – assign varying normative weight to such maxims. This, as should be clear, opens up wide opportunities for the appropriation of dutiful maxims for the actual elevation of the particular.

Turning now to Atterton’s contention that our own happiness is subsidiary to other imperfect duties (*viz.*, moral perfection and beneficence), his argument begins with the recognition that there must be a limit to beneficence, otherwise one would compromise his own happiness and also bring about his own poverty and need for beneficence. Thus,

unlimited beneficence is “self-defeating” (Atterton, 2007, 151). In response, Atterton identifies two passages in Kant which provides a limit to beneficence. The first is the casuistical question considered above which asserts that one ought not to be beneficent to the point of needing beneficence. Second is the more germane passage where Kant contends that sacrificing one’s own happiness in realizing one’s duty of beneficence is contradictory: “For a maxim of promoting others’ happiness at the sacrifice of one’s own happiness, one’s true needs, would conflict with itself were it made a universal law” (*MM*, 6:393). Atterton takes these passages to show that there are indeed limits to beneficence, but that these limits remain unclear. To discover these limits, he turns first to Kant’s discussion regarding how promoting one’s own happiness can never be a duty, and second to Kant’s insistence that our happiness not be neglected.

This maneuver is too quick. Atterton immediately turns from Kant’s insistence that sacrificing one’s happiness for the happiness of others would generate a contradiction and, thus, would constitute a violation of a negative duty to a consideration of Kant’s argument for why we do not have a duty to *promote* our happiness and only an indirect duty not to *neglect* our happiness. This abrupt transition seems to be based on an elision of a duty to *promote* and a duty to *ensure* one’s happiness, and Atterton’s quick analysis leads to a severe limitation of duty concerning our own happiness.

Atterton is right to interpret Kant as contending that we cannot have a duty to promote our own happiness. He also correctly recognizes Kant’s argument for an indirect duty not to *neglect* our happiness: neglecting our happiness and prosperity presents a temptation to vice and, thus, endangers our moral integrity. From these premises, however, Atterton argues that we can *only* have an indirect duty not to neglect our

happiness; that we must maintain a certain level of happiness as a *means* to preserving our moral integrity only (Atterton, 2007, 153). In this way, our own happiness is not directly an object of our duty, but only a consideration subsidiary to our moral perfection and beneficence. “The pursuit of my happiness is only ancillary to the fulfillment of imperfect duties; it is ‘indirectly’ a duty in the sense that it is a non-moral means to an obligatory end that lies outside it” (Atterton, 2007, 153).

This deduction is mistaken. Though Kant certainly does argue that we have an indirect duty to maintain our happiness as a means to maintaining moral integrity, this indirect duty is not our only duty concerning our happiness. We also have a strict duty to *ensure equal consideration* of our own happiness.

Returning to the above consideration of sacrificing our happiness for the sake of other's happiness, Kant argues that making one's own happiness subsidiary to the happiness of others would generate a contradiction using the universalization test. This contradiction is easy to see: If everyone were to make their happiness subsidiary to everyone else's, this would entail a constant lowering of one's happiness by everyone such that no one's happiness is equal to or higher than another's; however, this is inconceivable insofar as at least one person's happiness must be elevated in order for other's to be lower. Interestingly, this generation of a contradiction is analogous to the contradiction generated by the slavery maxim: if everyone abrogates their autonomy, then there is no one left for whom one can be a slave. Similarly, if everyone makes their happiness subsidiary, then there is no one left relative to whom my happiness is subsidiary. What this entails is that I have a duty to maintain an equal interest in my own happiness. Indeed, since the maxim of making my happiness subsidiary to the happiness

of others fails the contradiction in conception test, I have a strict duty to give my happiness the same consideration I give the happiness of others.

This equality and universal consideration of happiness accords with Kant's core theoretical commitment, and is expressly supported by him in terms of the Formula of Humanity. Kant holds that every rational being is an end-in-itself, and as such, from the standpoint of the moral law, no rational being's happiness could *not be* due equal consideration even by the agent himself. Kant expresses this perspective of the moral law as follows: "But since all *others* with the exception of myself would not be *all*, so that the maxim would not have within it the universality of a law...the law making benevolence a duty will include myself, as an object of benevolence, in the command of practical reason" (*MM*, 6:451). Since I must abide by the moral law, it follows directly that my happiness must be an object of benevolence, just as is everyone else's. Kant continues and explains that this does not entail that "I am thereby under obligation to love myself" since we are all (according to Kant) naturally and unavoidably inclined to our own happiness. It is important to have a clear understanding of what this means: I cannot be obligated to actively love myself, since active self-love comes naturally and one is obligated reluctantly. This impossibility of obligation, however, does not imply that my happiness is not due – even by me – equal consideration. And it does mean that I am *permitted* to love myself – that is to seek my own happiness. Crucially, this permission is not given as a means to some other duty. And though this language of "permittance" may sound weak (as if there is no duty involved), Kant's preceding contextualization of ensuring my own happiness in terms of the universalization test makes clear the necessity of such permission. That is, my own happiness *must* be permitted for its own sake lest I

impermissibly elevate the happiness of others above my own. If I were to *restrict* the realization of my happiness such that another's happiness was given priority, I would be acting contrary to duty.

This duty of permitting or ensuring my own happiness has significant practical implications. These can be seen by considering the conditions Atterton enumerates for when promoting my own happiness is permitted. There are two such conditions:

(a) promoting one's own happiness on those occasions when one can do nothing to fulfill an imperfect duty, and (b) promoting one's own happiness when it is a permissible means to fulfilling of an imperfect duty. (Atterton, 2007, 146)

These conditions make it so that one may permissibly pursue his own happiness only *after* one has considered the happiness of others: if I cannot be beneficent, then I may promote my happiness; and if my happiness is coextensive with beneficence, then I may (concurrently) promote my happiness. Abiding by either of these conditions, however, is impermissible insofar as each entail giving a lexical priority to beneficence (i.e., the happiness of others), and such a lexical priority requires not giving my own happiness equal consideration. As such, I am required to – as it were – consider my own happiness “side by side” with the happiness of others. Accordingly, I may pursue my own happiness even when I *can* do something to fulfill the duty of beneficence and when promoting my own happiness does *not* coincide with promoting another's happiness. Consequently, Atterton's contention about what R does prohibit is false: “What R does rule out is (c) making oneself an exception to the performance imperfect duties merely in the interests of inclination” (*ibid.*). If I were to never make an exception to a maxim of beneficence for my own interests and inclinations, then I would always be giving the happiness of others priority over my own happiness. To always give the happiness of others priority over my

own happiness violates my strict duty to give equal consideration to my own happiness. What I cannot do is always give priority to my own happiness (i.e., adopt a maxim of indifference) or fail to act beneficently when my own happiness is compatible with beneficence (e.g., if I fail to walk across a room to heal someone or save a baby from drowning in puddle or any number of other actions that require nominal effort and do not infringe on promoting my own happiness).

This analysis, I think, accords with Kant's repeated and emphatic insistence that no specific limit may be determined regarding the extent of beneficent action, and that each agent must determine for himself: "How far [the duty of beneficence] should extend depends, in large part, on what each person's true needs are in view of his sensibilities, and it must be left for each to decide for himself" (*MM*, 6:393; the same view is repeated in the following paragraph – subsection b – at 6:394). In light of the foregoing, this discretion cannot be construed as compatible with the subordination of one's happiness to the happiness of others.

Lastly, the above passage which clarifies the necessity of equal consideration of happiness in terms of the universalization test is compounded by a fascinating passage where Kant suggests that active, practical beneficence may vary in degree proportionate to those closer to me whom I love. That is, contrary to critics who argue that Kantianism is opposed to our strongly held intuitions of partiality,<sup>29</sup> we may act in a way that privileges those closest to us, including ourselves, so long as we are not *indifferent* to the happiness of others. Kant writes:

Now the benevolence present in love for all human beings is indeed the greatest in its *extent*, but the smallest in *degree*; and when I say that I take an interest in this

---

<sup>29</sup> See, for example, Bernard Williams (1982b) and Simon Keller (2013).



human being's well-being only out of my love for all human being, the interest I take is as slight as an interest can be. I am only not indifferent with regard to him. (*MM*, 6:451)

One might expect Kant to immediately condemn such variance in degrees of benevolence, and argue that we must eschew beneficence construed as the mere absence of indifference, or to argue that such "benevolence" is not equivalent to the duty of beneficence, and then argue for its replacement with universal and equal beneficence. Kant, however, argues that the fact that some are closer to me than others (including the closeness of myself) implies that such universal and equal beneficence would be contradictory:

Yet one human being is closer to me than another, and in benevolence I am closest to myself....If one is closer to me than another (in the duty of benevolence) and I am therefore under obligation to greater benevolence to one than to the other but am admittedly closer to myself (even in accordance with duty) than to any other, then it would seem that I cannot, without contradicting myself, say that I ought to love every human being as myself, since the measure of self-love would allow for no difference in degree. (*ibid.*)

First, Kant clarifies that he is considering the "duty of benevolence," which he occasionally uses as a substitute for "beneficence." Moreover, he argues that equal beneficence involves a contradiction of prescriptions. On the one hand, because I am closer to some than to others, as a matter of duty I am obligated to have greater benevolence to those to whom I am close (such as my parents). This, apparently, includes myself insofar as I am closest to myself. On the other hand, if I accept the commandment to love others as myself, then this commandment would allow for "no difference in degree" which directly contradicts my obligation to show a greater degree of benevolence to those closest to me. Kant, not wanting to abandon the commandment of universal and equal love (love your neighbor as yourself), resolves this conflict by distinguishing

between “benevolence in *wishes*,” which can be universally and equally extended, and *active* benevolence, which may vary in degree: “For in wishing I can be *equally* benevolent to everyone, whereas in acting I can, without violating the universality of the maxim, vary the degree greatly in accordance with the different objects of my love (one of whom concerns me more closely than another)” (*MM*, 6:452, emphasis in underlining is mine).

Thus, the universalization test both requires that I ensure my own happiness, and is compatible with *privileging in action* my own happiness, and the happiness of those closest to me, over the happiness of others. What the duty of beneficence requires is that I am not indifferent to the happiness of others, and that because of this universal and equal interest, I use the means available to me to promote the happiness of others so far as possible. This last qualification includes the limits established by negative duties, which includes the proscription against not giving equal consideration to my own happiness. How, when, and to what degree I realize this duty remains a matter of individual discretion (see, again, *MM*, 6:393-394).

What are we to conclude from the foregoing? First, that the manifold of duties is not circumscribed. In addition to the duty of beneficence, I have genuine and independent positive duties to myself (*viz.*, the duty of natural perfection) as well as a strict duty to ensure equal consideration of my own happiness. Accordingly, I may pursue my own interests and inclinations even when such pursuit would preclude particular actions of beneficence. Moreover, I am permitted to privilege the happiness of those closest to me, including myself, over the happiness of others so long as I do not become practically indifferent to their happiness.

Of greatest import, these conclusions have clear implications for the success of Hegel's objection. As Atterton recognized, if we have a duty regarding our own happiness then exceptions may be made to our duty of beneficence grounded in our own inclinations and interests. Also, we have a duty to pursue our own natural perfection. Lastly, in our actions we may privilege our own happiness over the happiness of others close to us. All of these duties may be appropriated in the actual service and elevation of the particular over the universal. In other words, the formal latitude inextricable from Kant's theory of positive duties enables the individual subject to find something *positive* and conformable to duty and, thus, to realize evil.

#### IV. CONCLUSION

This chapter presents three arguments. In two of these arguments Hegel's subjectivism objection is central. First, I adopted and explicated a canonical objection concerning the emptiness of the moral will and the unavoidability of subjectivism. This objection on its own was found wanting, and so I supported it by appropriating and applying arguments from Hegel's *Phenomenology* concerning the metaphysical and theoretical indefensibility of moral purity. Second, I argued that the limitation of actual negative duties argued for in chapter 4 extends the scope of possible positive duties. While brief, this argument is critical as it shows that the Formula of Humanity cannot proscribe traditionally opprobrious actions such as euthanasia and conventionally accepted deception, and that this line of reasoning may be extended to other actions. Lastly, I argued that Hegel's subjectivism objection (which is, per the arguments in the preceding section, unavoidable) in conjunction with the latitude of Morality regarding positive duties entails

the capacity for the subject to elevate the particular above the universal, and thus to actualize evil.

Each of these arguments contributes to recognizing the need for the sublation and comprehension [*aufgehoben*] of Morality and the transition to Ethical Life. Ethical Life provides content-full framework and resources, in the form of historically, culturally, and discursively contingent content, necessary for determining particular duties. One may object that by substituting contingent content for subjective determinations we are trading one relativism for another. However, as has been emphasized through this dissertation, the transition to Ethical Life does not entail the incorporation of *any* contingent, content-full framework; but rather, a standpoint that incorporates Morality and through the State actualizes freedom.

## Conclusion

The animating contention of this dissertation is that Hegel's emptiness critique remains relevant for Kantian moral philosophy. The significance of Hegel's critique is both broad and severe. However, the relevance of Hegel's emptiness critique has been challenged on two fronts: first, that Hegel's emptiness objection misunderstands Kant's position regarding the Formula of Universal Law and the functional role of the universalization procedure; second, that Hegel's objection is limited to the Formula of Universal Law, and the content-full Formula of Humanity undermines any charges of general emptiness. Both of these challenges are erroneous, and the specific objective of the dissertation is to show how.

The dissertation proceeds in three parts corresponding to these rebuttals: first, in chapter 1 I provide an alternate interpretation of Hegel's emptiness objection that contextualizes the objection, thereby showing how it is an element of a larger emptiness critique; second, in chapters 2 and 3 I consider and defend the applicability and severity of the emptiness critique for the Formula of Universal Law; lastly, in chapters 4 and 5 I argue that Hegel's critique is not limited to the Formula of Universal Law, but can be extended to the Formula of Humanity, and is thus broadly relevant for Kantian moral philosophy. These arguments are all framed in terms of Hegel's general contention that Morality cannot furnish an immanent doctrine of duties unless empirical content is imported: when a determination or derivation of duty works, it works by importing illicit content; when illicit content is not imported, the determination or derivation of duty fails. Morality is empty because the duty it commands is abstract: that the moral agent must do *right* and promote the *good*. Morality can take us no further.

Though the dissertation has a narrow focus, the consequences for Kantian moral theory are severe and fundamental. I introduced the dissertation by framing the debate in terms of the pure, *a priori* foundations of Kant's moral philosophy. An *a priori* foundation is essential to Kant, and continues to be for Kantian moral philosophy, in two respects: First, a pure moral foundation provides a stable and unified sources for all subsequent and ancillary "laws and concepts"; this enables genuine morality to eschew the confusion of different and competing sources of morality. Second, a pure rational foundation sets morality above competing empirical and contingent sources of morality; in this way, the *a priori* foundation of Morality is the ground of its supremacy. The unity and supremacy of Morality justifies its claim of objectivity and universality.

Hegel's emptiness critique does not directly challenge the objectivity or universality of Morality (i.e., it does not claim that Morality does not have a ground for abstract objectivity and universality), but rather discloses the cost of a pure foundation. Morality gains objectivity and universality by forsaking *actuality*. In Hegel's terminology, Morality is one-sided and incomplete: Morality contains the concept of duty and categorically commands its realization; however, from the pure standpoint of Morality one's particular duty remains undetermined. This internal inconsistency, according to Hegel, must be resolved by incorporating abstract Morality into the more complete dialectical moment of Ethical Life, or *Sittlichkeit*. Ethical Life provides the particular historical, cultural, and discursive content necessary for determining particular duties. As a necessary condition of actual duty, however, *Sittlichkeit* is foundational to practical philosophy. *Sittlichkeit* is also empirical – historically, culturally, and discursively specific. Thus, if Hegel's arguments are effective, practical philosophy is

fundamentally contingent.

It is important to emphasize the radical quality of Hegel's critique. An important element of Kantian Morality – an element that has not always been properly appreciated – is the requisite practical knowledge for the application of formal or objective duty. Thus, for example, Kant emphasizes the need to practice beneficence with attention and care lest one cause greater harm through embarrassment, shame, or (we may add) unwise or counterproductive means (see *MM*, 6:453; see also Baron and Falsmy [2009] and Herman [2008, 36]). Barbara Herman has extensively advanced this idea of requisite contingent knowledge and argued for its essential role in practical judgment (see Herman, 1993). Indeed, the essential role of contingent content in practical judgment undergirds Herman's argument for the compatibility of Kantianism and a plurality of subcultures and local values (Herman, 2008, 33ff). The objectivity and universality of Kantianism, then, need not be seen as rule-bound or intolerant.

A Hegelian may agree that contingent practical knowledge is essential to the realization of actual duty; indeed, Hegel argues in the *Phenomenology* (§425) that a principle of beneficence – to actively love thy neighbor as thyself – is insufficient without practical intelligence. However, the emptiness critique as presented here does not concern the need for practical knowledge of possibly diverse social conditions and contexts in the application of objective duties. Rather, the emptiness critique is concerned with the role of contingent content in the determination or derivation of substantive duties – not being engaged in killing, deception, theft and similar actions. According to the emptiness critique, fundamental and controversial duties cannot be determined independently from contingent conventional, subjective, and normative content.

To bring out this contrast with practical knowledge, let us briefly reconsider select examples from the text. I argued in chapter 1 that the success and effectiveness of the universalization procedure depends on the antecedent incorporation of heteronomous content – specifically, the normative specification in terms of which a maxim is tested must already include the proscription determined by the universalization test. In this way, the universalization test does not fail to function; rather, its results are vacuous and its use is superfluous.

The consequence of this superfluity, which may initially seem trivial, is dire: a maxim of slavery, for example, formulated with ordinary gradational specificity (a gradational specificity similar to other undertakings in the world of business) is universalizable within a counter-factual *Sittlichkeit* where a proscription against slavery is absent. This is because the maxim of slavery does not unfairly take advantage of the antecedently determined conventional proscription against slavery – that is, the maxim is not exceptional in character. However, whether or not a maxim is exceptional in character is a product of the normative specification posited in the will and maxim of the agent and in terms of which the maxim is tested. In terms of the Formula of Universal Law, then, the permissibility or impermissibility of a particular and situational maxim of slavery is dependent on contingent, conventional norms.

I argued for similarly dire consequences in chapter 2. The same example of slavery and slave-holding, which was presented by Onora O'Neill as a blatantly impermissible maxim, was used to show the challenges of construing maxims as governing principles. I argued that, contrary to her own criterion of maxim formulation, O'Neill formulates the slavery and slave-holding maxims in terms of specific social and



economic roles. With such a high degree of gradational specificity, however, false-positives proliferate – virtually any career choice construed in terms of a specific social and economic position is non-universalizable and not justifiable.

In chapter 2 I also considered the problem of natural actions and Korsgaard's solution. Korsgaard argues that the maxims of natural actions must incorporate specific intentions which take into account security as a "special object of human interest." I argued, first, that this solution is susceptible to Hegel's subjectivism objection, which contends that the specific intentions incorporated into one's maxim, and which thereby determine what is essential to an action, consist in the subjectively determined good one sees in the action; second, I argued that while security may be a "special object of human interest" it need not be the particular end of an action, but only a necessary means or background requirement. A change in intention as well as the operation of security in one's maxim, however, allows for egregiously wrong actions (such as killing a colleague in order to secure a competitive position). The consequence of extending Hegel's emptiness critique (in the form of the subjectivism objection) to natural actions is not merely the identification of possible false-negatives, no matter how outrageous they may be; rather, the consequence of greater importance is that determining *all* maxims of natural actions as permissible or impermissible depends on the subjective determination of what is good and essential in an action. Therefore, whether an action is contrary to or accordant with duty is fundamentally a matter of heteronomous and contingent content.

In chapters 4 and 5 I argued for the extension of Hegel's emptiness critique to the Formula of Humanity. In chapter 4 I argued that the derivation of negative duties, and in particular, suicide and lying, from the Formula of Humanity depends on independent

moral content and judgment. The need for additional moral content is most clearly seen in the derivation of negative duties (or duties of omission) in the *Doctrine of Virtue* which take a syllogistic form and incorporates intermediate premises. I argued, as is evident from Kant's own arguments, that these intermediate premises require independent justification in the form of ancillary syllogisms.

The premises of these ancillary syllogisms, however, are highly contentious or untenable. The ancillary argument supporting the duty against lying, for example, maintains that acting against the natural purpose of language (which is itself controversial) necessarily entails a renunciation of one's personality. This moral judgment is not grounded directly in the Formula of Humanity, but rather the contingent assumption that acting against the natural purpose of language necessarily opposes the rational ends of persons. I argue that this is an assumption that can be rejected.

The derivation of negative duties from the *Groundwork* differs significantly from that in *The Doctrine of Duty*. While still taking a syllogistic form, the derivation incorporates a common intermediate premise that does not require ancillary justification: that the prohibited acts in question (suicide and deception) use the relevant agent (oneself or another) as a mere means and not as an end. Suicide *disregard* one's humanity, debasing it in comparison to other values. Deception coopts another's rational will: by deceiving another person it is impossible for him to freely adopt and incorporate into his own will the ends set by the liar. I reject the derivation of the duty against deception by undermining the sufficiency of the criterion provided. I do this by providing ordinary scenarios where the criterion of disregard construed as circumventing or coopting another's will is satisfied without the act of deception being objectionable. I suggest that

this disparity between objectionable and unobjectionable cases of deception is derived from conventional, normative expectation; that is, our socially and culturally contingent norms ground our judgment about acceptable or unacceptable deception.

As a final example, in chapter 5 I extend Hegel's emptiness critique to the determination of positive imperfect duties in *The Doctrine of Virtue*. This extension presents the greatest challenge insofar as Kant does not derive or determine particular duties, but only general duties of self-perfection and practical love (beneficence) which are both plausible and immediately derivable from the Formula of Humanity. However, the generality of derived duties does not prevent applying Hegel's central contention that Morality cannot furnish an immanent doctrine of duty. I argue for an extension of Hegel's critique, first, by contending that the arguments limiting our negative duties in chapter 4 immediately entail a broadening of the scope of possible positive duties; for example, if euthanasia, for example, is permissible, then it *may be* our duty to assist in another's realization of this subjective end.

Second, from the standpoint of Morality, Hegel argues that the agent is confronted with a "manifold of duties" and which duty among the various possible duties is our actual duty remains undetermined. Hegel argues that this indeterminacy of one's actual duty, or the absence of "all valid determinations," entails the formal compatibility of "elevating the particular over the universal." That is, by subjectivizing the determination of actual duty, one may elevate his own interests and happiness above the promotion of universal welfare in a manner compatible with abstract duty. I elaborate Hegel's argument, and argue that we may understand Hegel as contending that where refraining from beneficence becomes impermissible or even opprobrious is left undetermined by

Morality. This implies that the subject determines when and if and to what degree he has acted morally or immorally, dutifully or contrary to duty. If this is true, then the subject may *in practice* elevate the particular above the universal while remaining formally in accord with duty. This, in turn, implies something truly radical: by choosing among goods, one may act in accordance with formal Morality while doing *evil*.

The foregoing, of course, is only a summary of a selection of arguments. This summary serves the purpose, however, of emphasizing the radical nature and extensive scope of Hegel's emptiness critique, the relevance and severity of which I have argued for here.

## Bibliography

- Acton, H.B. 1975. "Introduction." In *Natural Law: The Scientific Ways of Treating Natural Law, Its Place in Moral Philosophy, and Its Relation to the Positive Sciences of Law*, First edition. Philadelphia: University of Pennsylvania Press.
- Allison, Henry E. 1990. *Kant's Theory of Freedom*. Cambridge University Press.
- Ameriks, Karl. 2000. "The Hegelian Critique of Kantian Morality." In *Kant and the Fate of Autonomy: Problems in the Appropriation of the Critical Philosophy*, 309–37. Cambridge, U.K. ; New York: Cambridge University Press.
- Anscombe, G. E. M. 1958. "Modern Moral Philosophy." *Philosophy* 33 (124): 1–19. doi:10.1017/S0031819100037943.
- Aquinas, Thomas. 2013. *Summa Theologica, Vol. 2*. New York, NY: Cosimo Classics.
- Atterton, Peter. 2007. "A Duty to Be Charitable? A Rigoristic Reading of Kant." *Kant Studien* 98 (2): 135–155. doi:10.1515/KANT.2007.007.
- Aune, Bruce. 1980. *Kant's Theory of Morals*. Princeton, N.J: Princeton University Press.
- Balzan, Jurgen. 2016. "Updated: Parliament to discuss euthanasia before summer recess." *MaltaToday* [Online]. Available: [http://www.maltatoday.com.mt/printversion/66641/#.WPn4u\\_nyu00](http://www.maltatoday.com.mt/printversion/66641/#.WPn4u_nyu00) (accessed April 21, 2017).
- Baron, Marcia W. 1998. "Love and Respect in the Doctrine of Virtue." *The Southern Journal of Philosophy* 36 (S1): 29–44. doi:10.1111/j.2041-6962.1998.tb01776.x.
- Baron, Marcia W., and Melissa Seymour Fahmy. 2009. "Beneficence and Other Duties of Love in the Metaphysics of Morals." In *The Blackwell Guide to Kant's Ethics*, edited by Thomas Hill, First edition, 211–28. Chichester, U.K. ; Malden, MA: Wiley-Blackwell.
- Beiser, Frederick C. 2009. "'Morality' in Hegel's Phenomenology of Spirit." In *The Blackwell Guide to Hegel's Phenomenology of Spirit*, edited by Kenneth R. Westphal, First edition. Malden, MA: Wiley-Blackwell.
- Bennett, Jonathan. 1974. "The Conscience of Huckleberry Finn." *Philosophy* 49 (188): 123–34.
- Brown, Chris. 2008. "Kant and Therapeutic Privilege." *Journal of Medicine and Philosophy* 33 (4): 321–336.
- Bykova, Marina. 2015. "Will and Freedom." In *The Bloomsbury Companion to Hegel*, edited by Allegra de Laurentiis and Jeffrey Edwards, 265–69. New York: Bloomsbury Academic.
- Caygill, Howard. 1995. *A Kant Dictionary*. First edition. Malden, MA: Wiley-Blackwell.
- Cholbi, Michael. 2014. "Kant on Euthanasia and the Duty to Die: Clearing the Air." *Journal of Medical Ethics*, September, medethics-2013-101781. doi:10.1136/medethics-2013-101781.
- Coleman, Eli. 2003. "Masturbation as a Means of Achieving Sexual Health." *Journal of Psychology & Human Sexuality* 14 (2–3): 5–16.
- Denis, Lara and Wilson, Eric, "Kant and Hume on Morality", *The Stanford Encyclopedia of Philosophy* (Fall 2016 edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2016/entries/kant-hume-morality/>.

- Dietrichson, Paul. 1990. "Kant's Criteria of Universalizability." In *Foundations of the Metaphysics of Morals and Critical Essays*, edited by Robert Paul Wolff, translated by Lewis White Beck, First edition, 163–207. New York; London: MacMillan Pub Co.
- Donagan, Alan. 1979. *The Theory of Morality*. Revised edition. Chicago: University of Chicago Press.
- Engelhardt, H. Tristram. 2000. *The Foundations of Christian Bioethics*. First edition. Lisse (Netherlands); Exton, PA: M & M Scrivener Press.
- Findlay, J. N. 1976. *Hegel: A Re-Examination*. New York: Oxford Univ Press.
- Formosa, Paul. 2014. "Dignity and Respect: How to Apply Kant's Formula of Humanity." *The Philosophical Forum* 45 (1): 49–68. doi:10.1111/phil.12026.
- Franco, Paul. 2002. *Hegel's Philosophy of Freedom*. New edition. New Haven, Conn.; London: Yale University Press.
- Freyenhagen, Fabian. 2012. "The Empty Formalism Objection Revisited: §135R and Recent Kantian Responses." In *Hegel's Philosophy of Right*, edited by Thom Brooks, First edition. Chichester, West Sussex: Wiley-Blackwell.
- Galvin, Richard. 1991. "Does Kant's Psychology of Morality Need Basic Revision?" *Mind* 100 (2): 221–236.
- Geiger, Ido. 2007. "Why Does Hegel Charge Kant's Moral Theory with Emptiness?" In *The Founding Act of Modern Ethical Life: Hegel's Critique of Kant's Moral and Political Philosophy*, First edition, 9–32. Stanford, Calif: Stanford University Press.
- Guevara, Daniel. 1999. "The Impossibility of Supererogation in Kant's Moral Theory." *Philosophical and Phenomenological Research*, 593–624.
- Guyer, Paul. 2014. *Kant*. Second edition. London New York: Routledge.
- Hardwig, John. 1983. "Action from Duty but Not in Accord with Duty." *Ethics* 93 (2): 283–90.
- Hegel, G. W. F. 1971. *Philosophy of Mind*. Translated by William Wallace and A. V. Miller. First edition. Oxford: Clarendon Press.
- . 1975. *Natural Law: The Scientific Ways of Treating Natural Law, Its Place in Moral Philosophy, and Its Relation to the Positive Sciences of Law*. Translated by T. M. Knox. First edition. Philadelphia: University of Pennsylvania Press.
- . 1977. *Phenomenology of Spirit*. Translated by A. V. Miller. Revised edition. Oxford: Oxford University Press.
- . 1988. *Faith and Knowledge*. Albany: State Univ of New York Press.
- . 1991a. *The Encyclopaedia Logic: Part I of the Encyclopaedia of the Philosophical Sciences with the Zusatz*. Translated by T. F. Geraets, W. A. Suchting, and H. S. Harris. Indianapolis: Hackett Publishing Company, Inc.
- . 1991b. *Elements of the Philosophy of Right*. Edited by Allen W. Wood. Translated by H. B. Nisbet. Revised edition. Cambridge: Cambridge University Press.
- . 1971. *Philosophy of Mind*. Translated by William Wallace and A. V. Miller. First edition. Oxford: Clarendon Press.
- Henson, Richard G. 1979. "What Kant Might Have Said: Moral Worth and the Overdetermination of Dutiful Action." *The Philosophical Review* 88 (1): 39. doi:10.2307/2184778.
- Herman, Barbara. 1993. "Leaving Deontology Behind." In *The Practice of Moral Judgment*, 208–42. Cambridge: Harvard University Press.

- . 1993a. "Moral Deliberation and the Derivation of Duties." In *The Practice of Moral Judgment*, 132–58. Cambridge: Harvard University Press.
- . 1993b. "Murder and Mayhem." In *The Practice of Moral Judgment*, 113–31. Cambridge: Harvard University Press.
- . 1993c. "Mutual Aid and Respect for Persons." In *The Practice of Moral Judgment*, 45–72. Cambridge: Harvard University Press.
- . 1993d. "On the Value of Acting from the Motive of Duty." In *The Practice of Moral Judgment*, 1–22. Cambridge (MA) etc.: Harvard University Press.
- . 1993e. "The Practice of Moral Judgment." In *The Practice of Moral Judgment*, 73–93. Cambridge: Harvard University Press.
- . 1998. "Making Room for Character." In *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*, edited by Stephen Engstrom and Jennifer Whiting. Cambridge: Cambridge University Press.
- . 2008. "Pluralism and the Community of Moral Judgement." In *Moral Literacy*, 29–50. Cambridge, Mass.: Harvard University Press.
- Hill, Thomas E. 1971. "Kant on Imperfect Duty and Supererogation." *Kant-Studien* 62: 55–76
- . 1991a. "Autonomy and Benevolent Lies." In *Autonomy and Self-Respect*, 25–42. Cambridge, England ; New York: Cambridge University Press.
- . 1991b. "Self-Regarding Suicide: A Modified Kantian View." In *Autonomy and Self-Respect*, 85–103. Cambridge England ; New York: Cambridge University Press.
- Hinman, Lawrence M. 1983. "On the Purity of Our Moral Motives: A Critique of Kant's Account of the Emotions and Acting for the Sake of Duty." *The Monist* 66 (2): 251–267.
- Hogarth, Harriet, and Roger Ingham. 2009. "Masturbation among Young Women and Associations with Sexual Health: An Exploratory Study." *Journal of Sex Research* 46 (6): 558–567.
- Hoy, David Couzens. 1981. "Hegel's Morals." *Dialogue* 20 (01): 84–102.
- . 1989. "Hegel's Critique of Kantian Morality." *History of Philosophy Quarterly* 6 (2): 207–232.
- . 2009. "The Ethics of Freedom: Hegel on Reason as Law-Giving and Law-Testing." In *The Blackwell Guide to Hegel's Phenomenology of Spirit*, edited by Kenneth R. Westphal, First edition. Malden, MA: Wiley-Blackwell.
- Inwood, Michael. 1992. *A Hegel Dictionary*. First edition. Oxford, OX, UK: Wiley-Blackwell.
- Jehl, Douglas. 1994. "Surgeon General Forced to Resign by White House." *New York Times* [Online]. Available: <http://www.nytimes.com/1994/12/10/us/surgeon-general-forced-to-resign-by-white-house.html> (accessed April 21, 2017).
- Johnson, Robert and Cureton, Adam, "Kant's Moral Philosophy", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/spr2017/entries/kant-moral/>>.
- Kant, Immanuel. 1787. *Lectures on Ethics by Immanuel Kant*. Hackett Pub Co Inc.
- . 1999a. *Religion within the Boundaries of Mere Reason: And Other Writings*. Translated by Allen Wood and George Di Giovanni. Cambridge ; New York: Cambridge University Press.
- . 1999b. *Critique of Pure Reason*. Edited by Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.



- . 1999c. "Critique of Practical Reason." In *Practical Philosophy*, edited by Mary J. Gregor, New edition, 139–271. Princeton, N.J.: Cambridge University Press.
- . 1999d. "Groundwork of the Metaphysics of Morals." In *Practical Philosophy*, edited by Mary J. Gregor and Jens Timmermann, New edition, 37–108. Princeton, N.J.: Cambridge University Press.
- . 1999e. "On a Supposed Right to Lie from Philanthropy." In *Practical Philosophy*, edited by Mary J. Gregor, New edition, 605–15. Princeton, N.J.: Cambridge University Press.
- . 2011. "Conjectural Beginning of Human History." In *Anthropology, History, and Education*, translated by Allen Wood, Reprint edition, 160–75. Cambridge: Cambridge University Press.
- Kant, Immanuel. 1999. "The Metaphysics of Morals." In *Practical Philosophy*, edited by Mary J. Gregor, New edition, 353–59. Princeton, N.J.: Cambridge University Press.
- Keller, Simon. 2013. "Special Relationships and Special Reasons." In *Partiality*, 1–30. Princeton: Princeton University Press.
- Korsgaard, Christine M. 1996a. *The Sources of Normativity*. Edited by Onora O'Neill. Cambridge; New York: Cambridge University Press.
- . 1996b. *Creating the Kingdom of Ends*. Cambridge University Press.
- Leighton, Alexander H., and Charles C. Hughes. 1955. "Notes on Eskimo Patterns of Suicide." *Southwestern Journal of Anthropology* 11 (4): 327–338.
- Lo, Ping-cheung. 1981. "A Critical Reevaluation of the Alleged 'Empty Formalism' of Kantian Ethics." *Ethics* 91 (2): 181–201.
- . 1987. "Treating Persons as Ends."
- Lottenbach, Hans, and Sergio Tenenbaum. 1995. "Hegel's Critique of Kant in the Philosophy of Right." *Kant-Studien* 86 (2): 211–230.
- Mahon, James Edwin. 2006. "Kant and the Perfect Duty to Others not to Lie." *British Journal for the History of Philosophy* 14: 653–685.
- . 2009. "The Truth about Kant on Lies." *The Philosophy of Deception*, 201–224.
- McIntyre, Alison, "Doctrine of Double Effect", *The Stanford Encyclopedia of Philosophy* (Winter 2014 edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2014/entries/double-effect/>>.
- O'Neill, Onora. 1990a. "Consistency in Action." In *Constructions of Reason: Explorations of Kant's Practical Philosophy*, 81–104. Cambridge England ; New York: Cambridge University Press.
- . 1990b. "Reason and Autonomy in Grundlegung III." In *Constructions of Reason: Explorations of Kant's Practical Philosophy*, 51–65. Cambridge England ; New York: Cambridge University Press.
- . 1990c. "Universal Laws and Ends-in-Themselves." In *Constructions of Reason: Explorations of Kant's Practical Philosophy*, 126–44. Cambridge England ; New York: Cambridge University Press.
- . 2014. *Acting on Principle: An Essay on Kantian Ethics*. Second edition. Cambridge University Press.
- Pinkard, Terry. 2013. *Hegel's Naturalism: Mind, Nature, and the Final Ends of Life*. Reprint edition. New York etc.: Oxford University Press.
- Pippin, Robert B. 1995. "Hegel on the Rationality and Priority of Ethical Life." *Neue Hefte Für Philosophie*, no. 35: 95–126.



- . 2008. *Hegel's Practical Philosophy: Rational Agency as Ethical Life*. First edition. Cambridge, UK ; New York: Cambridge University Press.
- Rachels, James. 1999. *The Elements of Moral Philosophy*. McGraw-Hill College.
- Reath, Andrews. 2006. "Kant's Theory of Moral Sensibility: Respect for the Moral Law and the Influence of Inclination." In *Agency and Autonomy in Kant's Moral Theory: Selected Essays*, First edition, 8–32. Oxford : Oxford ; New York: Clarendon Press.
- Schroer, Jeanine Weekes, and Alexandra Bradner. 2011. "Experiment Month Results on Thomson's Violinist." *Experimental Philosophy*. July.  
<http://philosophycommons.typepad.com/xphi/2011/07/experiment-month-results-on-thomsons-violinist.html>.
- Sedgwick, Sally. 1988. "On the Relation of Pure Reason to Content: A Reply to Hegel's Critique of Formalism in Kant's Ethics." *Philosophy and Phenomenological Research* 49 (1): 59–80.
- . 2011. "'Letting the Phenomena In': On How Herman's Kantianism Does and Does Not Answer the Empty Formalism Critique." *Kantian Review* 16 (1): 33–47.  
doi:10.1017/S1369415410000063.
- . 2015. "Hegel on the Empty Formalism of Kant's Categorical Imperative." In *A Companion to Hegel*, edited by Stephen Houlgate and Michael Baur, First edition. Wiley-Blackwell.
- Stern, Robert. 2012. "On Hegel's Critique of Kant's Ethics: Beyond the Empty Formalism Objection." In *Hegel's Philosophy of Right*, edited by Thom Brooks, First edition. Chichester, West Sussex: Wiley-Blackwell.
- Thomson, Judith Jarvis. 1971. "A Defense of Abortion." *Philosophy and Public Affairs* 1 (1): 47–66.
- Timmermann, Jens. 2013. "Kantian Dilemmas? Moral Conflict in Kant's Ethical Theory." *Archiv Für Geschichte Der Philosophie* 95 (1): 36–64.
- Urmson, J. O. 1958. "Saints and Heroes." In *Essays in Moral Philosophy*, edited by A. I. Melden. University of Washington Press.
- Wallace, William, and A. V. Miller, trans. 1971. *Hegel's Philosophy of Mind*. First edition. Oxford: Clarendon Press.
- Westphal, Kenneth. 1991. "Hegel's Critique of Kant's Moral World View." *Philosophical Topics* 19 (2): 133–176.
- . 1993. "'The Basic Context and Structure of Hegel's Philosophy of Right'." In *The Cambridge Companion to Hegel*, edited by F. C. Beiser. Cambridge: Cambridge University Press.
- . 1995. "How Full Is Kant's Categorical Imperative." *Jahrbuch Für Recht Und Ethik/Annual Review of Law and Ethics* 3: 465.
- . 2015. "The Phenomenology of Spirit of 1807: Conspectus." In *The Bloomsbury Companion to Hegel*, edited by Allegra de Laurentiis and Jeffrey Edwards, 39–52. New York: Bloomsbury Academic.
- Will, George. 2015. "Is it Ethical to Affirm a Right to Die?" *The National Review* [Online]. Available: <http://www.nationalreview.com/article/423253/it-ethical-affirm-right-die-george-will> (accessed April 21, 2017).
- Williams, Bernard. 1982a. "Moral Luck." In *Moral Luck*, 20–39. Cambridge: Cambridge University Press.

- . 1982b. "Persons, Character and Morality." In *Moral Luck*, 1–19. Cambridge: Cambridge University Press.
- . 1982c. "Rawls and Pascal's Wager." In *Moral Luck*, 94–100. Cambridge: Cambridge University Press.
- Wolf, Robert Paul. 1974. *The Autonomy of Reason*. New York: Harper and Rowe. [Kindle edition].
- Wood, Allen. 1989. "The Emptiness of the Moral Will." *The Monist* 72 (3): 454–483.
- . 1990. *Hegel's Ethical Thought*. Cambridge England ; New York: Cambridge University Press.
- . 1999. *Kant's Ethical Thought*. Cambridge University Press.
- . 2007. *Kantian Ethics*. First edition. Cambridge ; New York: Cambridge University Press.
- . 2009. "Duties to Oneself, Duties of Respect to Others." In *The Blackwell Guide to Kant's Ethics*, edited by Thomas Hill, First edition, 229–51. Chichester, U.K. ; Malden, MA: Wiley-Blackwell.